

## ALIGNMENTS





QY	1021	ACAAGAAGATTTGTTAGTAACTGTGAGGACACAAATCAGAGGTTTGTGACTGAAGTG	1080
DB	1021	ACAAGAAGATTTGTTAGTAACTGTGAGGACACAAATCAGAGGTTTGTGACTGAAGTG	1080
QY	1081	GTCAATATCAATACCGTAGTTGGGATCTGGAGGAAATGACACAAATTTGTTTGGAA	1140
DB	1081	GTCAATATCAATACCGTAGTTGGGATCTGGAGGAAATGACACAAATTTGTTTGGAA	1140
QY	1141	TCAACATCAACCCAGCTGCTCTCAAGTACTGTGAGGACACACAGAGCTGTTAAAGTAT	1200
DB	1141	TCAACATCAACCCAGCTGCTCTCAAGTACTGTGAGGACACACAGAGCTGTTAAAGTAT	1200
QY	1201	TGCAATGCTCTCTCATCTTCATGAGTCTTCTGCACTCGAGGAGAACTGCAGATAGATG	1260
DB	1201	TGCAATGCTCTCTCATCTTCATGAGTCTTCTGCACTCGAGGAGAACTGCAGATAGATG	1260
QY	1261	TATTCGTTTTTGGAAATACAAACCACTCACACCTTACCTGATAGGACACTGGAATCA	1320
DB	1261	TATTCGTTTTTGGAAATACAAACCACTCACACCTTACCTGATAGGACACTGGAATCA	1320
QY	1321	GGTTTGCATCTGTCTGCTGCTCAAAATGCTCAACCACTAGTAAGCACACATGGTACTC	1380
DB	1321	GGTTTGCATCTGTCTGCTGCTCAAAATGCTCAACCACTAGTAAGCACACATGGTACTC	1380
QY	1381	CCAGAACCAAGATTTATTTGTTGAGATACCCCACTATGTCAGAGCTGGGACCTTTACCGG	1440
DB	1381	CCAGAACCAAGATTTATTTGTTGAGATACCCCACTATGTCAGAGCTGGGACCTTTACCGG	1440
QY	1441	CCATCTTATAGGTTCTCTATCTTCTGCACTCTCCAGATGACAGACTATTGTAACCTGG	1500
DB	1441	CCATCTTATAGGTTCTCTATCTTCTGCACTCTCCAGATGACAGACTATTGTAACCTGG	1500
QY	1501	AGCTGAGATGAAACGCTTAGGTTCTGGAATGTTTCCCTTCCCTTAATCAAGAAATAC	1560
DB	1501	AGCTGAGATGAAACGCTTAGGTTCTGGAATGTTTCCCTTCCCTTAATCAAGAAATAC	1560
QY	1561	TGAAGTGAATCGGAGCAATATCTCTGGAAGACTATCTATCAGTGAATGATCTCTGCG	1620
DB	1561	TGAAGTGAATCGGAGCAATATCTCTGGAAGACTATCTATCAGTGAATGATCTCTGCG	1620
QY	1621	GTTCGAGCCCAATCATGTGCGCATATTTCTAAGTTTGGGTTGCTGTGTAGAACTAAATTC	1680
DB	1621	GTTCGAGCCCAATCATGTGCGCATATTTCTAAGTTTGGGTTGCTGTGTAGAACTAAATTC	1680
QY	1681	TGAGCGGAGAACCACTAGTGGAAAAACCTTGAATATAAACAACCAACCAAGTAGCATC	1740
DB	1681	TGAGCGGAGAACCACTAGTGGAAAAACCTTGAATATAAACAACCAACCAAGTAGCATC	1740
QY	1741	TTTACCAACTGGGAGAGCCTTGGAGGAGCTATAAAGTTTTCATATGCTGCGCGTGAT	1800
DB	1741	TTTACCAACTGGGAGAGCCTTGGAGGAGCTATAAAGTTTTCATATGCTGCGCGTGAT	1800
QY	1801	ATTCTCGATCATGTAGTCTCATTTTATATGAAAGATGATACAAATGGGTAAT	1860
DB	1801	ATTCTCGATCATGTAGTCTCATTTTATATGAAAGATGATACAAATGGGTAAT	1860
QY	1861	TATTGCTCTGACCTATACATGCACTGAGTGTAGCAAGTTTATTTTATCTCTT	1920
DB	1861	TATTGCTCTGACCTATACATGCACTGAGTGTAGCAAGTTTATTTTATCTCTT	1920
QY	1921	TTTTTCT	1980
DB	1921	TTTTTCT	1980
QY	1981	TAAACAGAG 1988	
DB	1981	TAAACAGAG 1988	

RESULT 3  
AY128834  
LOCUS  
1541 bp mRNA linear PLN 07-AUG-2002

DEFINITION	Arabidopsis thaliana putative fizzy-related protein (At4g22910)
ACCESSION	AY128834
VERSION	AY128834.1
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 1541) Tripp, M., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M., Palm, C., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S., Narusaka, M., Pham, P., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUL-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu  The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Tripp, M., Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M., Kim, C., Lin, J., Liu, S., Pham, P., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.  Tripp, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.
FEATURES	Location/Qualifiers 1..1541 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="4" /clones="U24687" /note="This clone is in pENTR/SD-dtopo This is a cloned PCR product using RIKEN clone RAPL09-63-A14 (AY099585) as a template ecotype: Columbia" 1..1541 /gene="At4g22910" /note="synonym: F7H1.9.90" 1..1452 /gene="At4g22910" /codon_start=1 /product="putative fizzy-related protein" /protein_id="AA051234.1" /translation="MEEDEPTASNVITNSSSSNRNLSPAMNTPVVSLEIRNLINA NQSQSPSSLSRSIYSDRTPSRSGNFALFDLSPSKDGKDGAGSVATLURAM FGPTPEKRDITGFSRSRNPFRKTEHRSLSNFSFGVDDDSFGVSGSPVKAPRY PRSPYKVLADPALODDFVLNVDVAQNLVAGLNCVFLWNACSKVTKLCLDGAED SVCVSGWALRGTHTLAVGTSGKVOIPDASRCKRTWEGHRLRVGLAWGSSVLSGSD KLSLQKRDIRCQSDHVSKLACHSEVCGSLKNSVDNRELKSGGNDNRLFWYNOISSTOP VLKYSHTAAVKAIAWSPHVLGASGGTADRCIRFNNTTTHLSIDPCSOVCNL AWSKNVNLVSTHGYSQNIIVKIPYPTMSKIATLTGHTYRVLVLAVPDGGTITVGAG DETLLRFNVFPSPKSDTSSEIGSSFFGRTTIR"
gene	ecotype: Columbia 1..1541 /gene="At4g22910" /note="synonym: F7H1.9.90"
CDS	1..1452 /gene="At4g22910" /codon_start=1 /product="putative fizzy-related protein" /protein_id="AA051234.1" /translation="MEEDEPTASNVITNSSSSNRNLSPAMNTPVVSLEIRNLINA NQSQSPSSLSRSIYSDRTPSRSGNFALFDLSPSKDGKDGAGSVATLURAM FGPTPEKRDITGFSRSRNPFRKTEHRSLSNFSFGVDDDSFGVSGSPVKAPRY PRSPYKVLADPALODDFVLNVDVAQNLVAGLNCVFLWNACSKVTKLCLDGAED SVCVSGWALRGTHTLAVGTSGKVOIPDASRCKRTWEGHRLRVGLAWGSSVLSGSD KLSLQKRDIRCQSDHVSKLACHSEVCGSLKNSVDNRELKSGGNDNRLFWYNOISSTOP VLKYSHTAAVKAIAWSPHVLGASGGTADRCIRFNNTTTHLSIDPCSOVCNL AWSKNVNLVSTHGYSQNIIVKIPYPTMSKIATLTGHTYRVLVLAVPDGGTITVGAG DETLLRFNVFPSPKSDTSSEIGSSFFGRTTIR"
BASE COUNT	422 a 312 c 360 g 447 t
ORIGIN	
Query Match	35.9%; Score 719.8; DB 8; Length 1541;
Best Local Similarity	74.2%; Pred. No. 5.5e-156;



Matches	974;	Conservative	0;	Mismatches	312;	Indels	27;	Gaps	4;
QY	306	CCTCACCTTCTCGAACAATCTACTCCGATAGTTCATTCGAGTAGATCTGCTTCGAAAT	365						
Db	158	CATCACTATCAAGGCTCTATATACTCTGATAGATTTATCCCACTAGATCCGGATCCAAAT	217						
QY	366	TGCGTTTGTTCATATCAATATCTCCGAC-----AGAGGACGCGATGATAGTTCACGC	419						
Db	218	TGCGCTCTTTTCGATCTATCTCTCTTCCTAGTAAAGATGGTAAAGAGATGGAGCTGGCT	277						
QY	420	CTTATACGACTCTTCTCGAAGACGGCGTGTGTTTGGACCGGATGTTCCCGTCCGGTTACGC	479						
Db	278	CTTACGCTACTCTGTTGGTGGCGGATGTTGGTCTCGAG-----ACGC	322						
QY	480	CGGAAAAACCGACTCGCCGTGATGACATGCGGATAGGATATTTTAGGTATAGA	539						
Db	323	CGGAGAGAGAGATATTAATCGGTGTTTCTTCTTC-----AGGAATATTTTAGGTATAGA	379						
QY	540	CGGAGACGACAGTCCATGCACTCGCTTTTCGCGTTTATGGATGATGATTTTGTTCCTG	599						
Db	380	CGGAGACTCATCGGTCTTTGAAATCGTTTCTCCTTTTGGTGTGATGATGATCTCCTG	439						
QY	600	GTGTTAATCATATA--GTCCGGTTAAGGCTCTTAGGAGGTTCTCGATCGCCTTATAGG	656						
Db	440	GTGTTTCTCATAGTGTGCTCTGTTTAAAGCTCCCAAGGAAGTGGCGGATCGCGGTATAGG	499						
QY	657	TTTGTGATGCACCTGCTTTCGAAGATGATTTTATCTGAATCTGGTAGATTTGGTCTTCAAC	716						
Db	500	TATTGGATGCACCGCCTTCAGATGATTTTATCTGAATCTTGTGGATTTGGTCTGCAC	559						
QY	717	ACAATGTGTTGGCTGTGTTTCGGTAACTGTCTATCTCTGGAATGCTGTAGCAGCA	776						
Db	560	AAAATGTTCTAGCAGTGGGACTAGGGAACCTGTGTGATTTATGGAATGCTGTGAGCAGCA	619						
QY	777	AGGTAACTAAATATGTGATTTCCGGGTTCATGATTTGTTCTCTCTGTGGTTGGGCTC	836						
Db	620	AGGTACTAAGTTATGTGATCTCGAGCTGAGGATAGTGTGTTGCTCAGTGGGTTGGGCT	679						
QY	837	AACGTGTACTCATCTTGTCTGTTTGGAACTAACCAATGGTAAAGTTCAGATTTGGGATGCAAG	896						
Db	680	TACGTGAACTCATCTGCTGTTTGGAACTAGTACC CGGAAAGTTCAGATATGGGATGCGT	739						
QY	897	CAAGATCGAAGAGATTAAGATCAATGAGGGCCATCGGTTACGTGTCGGGCCCTCGGCT	956						
Db	740	CACCTGCAAGAGAACCAAGAACATGGAAGGTCACTGCTTAAGAGTTGGAGCCCTGGCAT	799						
QY	957	GGAGTTCATCTCTTGTCTCTGCTGCGAGGATAGAAATATTTATCAACGAGATATAC	1016						
Db	800	GGGTTTCAATCGGTTCTGTCTCTGTTAGTAAACTGTGACGACACAGAGTATTTCTTCAGAGACATAA	859						
QY	1017	GCAACAAGAGATTTTGTGTAGTAAACTGTGACGACACAAATCAGAGTTTGTGGACTGA	1076						
Db	860	GGTGTCAAGAGATCATGTCAATAATGGCAGGTCTAATACTGAGGTATGCGGACTCA	919						
QY	1077	AGTGGTCAATGATPAAACCGTCAGTTGGCATCTGAGGAAATGACAAACAAATTTGTTGTTT	1136						
Db	920	AGTGGTCTTATGACAACAGAGAGCTAGCACTGTTGGAAACGACATAGGCTTTTGTAT	979						
QY	1137	GGAATCAACACTCAACCCAGCGCTGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAG	1196						
Db	980	GGAACCAACATTTCAACAACAACCGGTTTTTGAATATAGTGAACAACATCTCAGCTGTGTTAAAG	1039						
QY	1197	CTATTGCAATGGTCTCCTCATCTTATGGAATTTCTTGTGATCTCGAGGAGGAACTGCAGATA	1256						
Db	1040	CCATTGCTTGTCTCCTCATGTTTATGAGGCTCTTGTCTTCTGTTGGTGTGTTACTGCTGATA	1099						
QY	1257	GATGTATTTGTTTTTGGAAATCAACCAACAACACTCACCTTTAGCTGTATGGACACTGGAA	1316						
Db	1100	GATGATACGTTTTTGGAAATCAACCAACGAATCTCAATTTAAGTTCCATAGATACTTGCA	1159						
QY	1317	GTCAAGTTTGCAAATCTTGTCTGGTCCAAATATGTCACGAACACTAGTAGACACATGGT	1376						
Db	1160	GTCAAGTATGCAATCTAGCTTGTCTAGAACGTAACGAGCTGTTTAGACACACGAT	1219						

```

/gene="At4g22910"
/codon_start=1
/product="putative fuzzy-related protein"
/protein_id="RAM20437.1"
/db_xref="GI:20466239"
/translation="MEEDPTASNVITNSNSMRNLSPAMNTPVVLBSRINRLINA
NOSQPSRSSLSRSIYSDFIPSRSGNFALFDLSPSPKDGKEDGAGSVATLLRAAM
FGPTEPERDITGSSSRNIFFKTEHRSNIFSPFVDDSDPGVSHSGPKAPKRV
PRSPYKLDALPDQDFYINLVDMAQNLAVGLNVCVILWNACSSKVKLADLGAED
SVCSVGMALRGTHLAVGSTKQVQIWDASRCRTRTMEGHLRVLGALWGLSVLSGS
RDKSLORDIRSCODHVSFKAGHSVEGLKWSYDNRELASGGDNRLFWNOHSQTP
VLKYSERTAAVKAJAWSPHVGLLASGGTADRCIRPWNTTNTLHSSLDTCQVCL
RMSKYNELVETHGVSQNIIVKPKPTMSKIATLTGTYAVLYLAVSPDQGIIVTGA
DETURFWNVFSPKSNQNDSEIGSFQRTIR"
BASE COUNT      544 a   374 c   393 g   554 t
ORIGIN

```

```

Query Match      35.9%; Score 719.6; DB 8; Length 1865;
Best Local Similarity 74.2%; Pred. No. 5.5e-158;
Matches 974; Conservative 0; Mismatches 312; Indels 27; Gaps 4;

QY 306 CCTCACTCTCGAACAATCTACTCCGATAGTTCATCCGAGTAGATCTCGTTCGAAAT 365
DB 224 CATCACTCAAGGCTCTATATCTCTGATGATTTATCCCGATAGATCCGATCCAAAT 283
QY 366 TCGCTTTGTTGATATCAATCTCCGAC-----AGAAGACGCGATGATGTTCCAGCG 419
DB 284 TCGCTTTTTCGATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 343
QY 420 CTATACGACTCTCTGAGACGGGTTGTTTGAACCGGATGTTGCCGTCCTCGGTTACGC 479
DB 344 CTTAGCTACTCTGTGTGGTGGGCGATGTTTGTGCTCTGAG-----ACGC 388
QY 480 CGGAATAAACCGACTCGCGCTGATGACATTCGCAATAGCAATATTTTAGGTATAAGA 539
DB 389 CGGAGAAGAGATATTAATCTCGGTTTCTCTCTCC---AGCAATATTTTAGGTATAAGA 445
QY 540 CGGAGACGAGACAGTCCATGACTCGCTTCGCGGTTATGAGATGATTTTGTCTCG 599
DB 446 CGGAGACTCATCGGCTTGAATTCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
QY 600 GTGTTAATCAT- --GTCCGGTTAAGGCTCTAGGAAGGTTCTTCGATCGCCTTATAAG 656
DB 506 GTGTTTCTCATAGTGTCTCTGTTAAAGCTCCAGAAAGTCCCGCATCGCGTATAAG 565
QY 657 TTTTGGTGCACCTCTTGAAGATGATTTTATCTGAATCTGATGATGATGATGATGAT 716
DB 566 TATTGATGACCGCGCTTGAAGATGATTTTATCTGAATCTGATGATGATGATGATGAT 625
QY 717 ACAATGTGTGCTTGTGTTGGTAACTGTGTCTATCTCTGGAATGCTTGTAGCAGCA 776
DB 626 AAAATGTTCTAGCAGTGGGACTAGGGAATCTGTGTATTTATGGAATGCTTGTAGCAGCA 685
QY 777 AGSTAACTAAATATGATGATTTGGGGTGTGATGATGATGATGATGATGATGATGATG 836
DB 686 AGTTACTAAGTATGATGATCTCGAGCTGAGGATGATGATGATGATGATGATGATGAT 745
QY 837 AACGTGGTACTCATCTCTGTTGGAACATAAATAGTGTAAAGTTTCAATTTGGATGCGAG 896
DB 746 TACGTGGAATCATCTGCTGTTGGAATAGTAGTACCGGAAAGTTTCAGATATGCGATGCG 805
QY 897 CAAGATGCAAGAGATAGATCAATGAGGCGCATCGTTACGTGTCGGGCTTGGCGCT 956
DB 806 CACGTGCAAGAGAACAGAACATGGAAGTCAATGCTTGAAGTTTGGAGCCCTGGCAT 865
QY 957 GGAGTTCACTCTTTTGTCTCTGTTGAGCGGATAAGATATTTATCAACGAGATATAC 1016
DB 866 GGGGTTCACTCGGTTCTGTCATCTGTTAGCAGACAGAGATATTTCTTCAGAGACATAA 925
QY 1017 GCACACAGAGATTTTCTTGTAGTAACTGTGACGACCAATATCGAGGTTTGTGGACTGA 1076
DB 926 GGTGTCAAGAGATCATGTCAGTAAATTTGGCAGTTCATAATCTGAGGTATCGGACTCA 985

```

```

QY 1077 AGTGTCTCATATGATAACCGGTGAGTTGGCATCTCGAGAAATGACAAATATGTTTGT 1136
DB 986 AGTGTCTTATGACAAAGAGAGCTAGCATCTGTTGGAACGACATAGCTTTTGTAT 1045
QY 1137 GGAATCAACACTCAACCCGCTGCTCTCAAGTACTGTGTAGGACACAGACGCTGTTAAAG 1196
DB 1046 GGAACCAACATTCACACCAACCGGTTTGAATATATAGTGAACACACTGCGAGCTGTTAAAG 1105
QY 1197 CTAATTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1256
DB 1106 CCAATTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1165
QY 1257 GATGATTTCTGTTTGTGAATCAACACCAACTCACACCTTAGCTGTATGACACTGGA 1316
DB 1166 GATGATACGTTTGTGAATCAACACCAACTCACTTAAAGTTTCCATAGATATCTGCA 1225
QY 1317 GTCAGGTTTCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1376
DB 1226 GTCAGGTTTCAATCTAGCTTGTCTTAAAGACGTTAAACGAGCTTGTAGCACACGAT 1285
QY 1377 ACTCCGACACCAAGATTTGTTTGGAGTACCCACTATGTCAAGCTGGGCTCTTA 1436
DB 1286 ACTCTCAGAACCAATCAITGTTCTGGAATACCCCAACCATGTCCAAATTTGCTACTTAA 1345
QY 1437 CCGGCCATCTTATPAGGTTTCTCTATCTTCCCATCTCTCCAGATGACAGACTTATTGTA 1496
DB 1346 CCGGTCAACATACCGAGTCTTATACCTTGGGTTTCAACCGATGACAGACGATTTGTA 1405
QY 1497 CTGAGCTGAGATGAAACGCTTAGGTTCTGGAATCTTCCCTTCCCTTAAATACACAG 1556
DB 1406 CAGGACGAGATGAAACCTTAAAGTTCTGGAATCTTCCCTTCCCTTAAATCTCAGA 1465
QY 1557 ATACTGAAGTGAATCGGAGCATTTCTCTCGAAGAACTACTACAGGTGA 1609
DB 1466 ACAGGATAGTGAATCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1518

```

```

RESULT 5
AY091235 1459 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana putative Srw1 protein (At4g11920) mRNA,
DEFINITION complete cds.
ACCESSION AY091235
VERSION AY091235.1 GI:20259634
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1459)
Yanada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.J., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1459)
Yanada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.J., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA

```

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki.M., Narusaka.M., Ishida.J., Sacou.M., Kamiya.A., Sakurai.T., Carninci.P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAFL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,I., Chen,H., Cheuk,P., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES  
source

Location/Qualifiers  
1. .1459

/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosomes="4"  
/clone="U18065"  
/note="This clone is in PUN1 51."  
eotype: Columbia  
1. .1459

gene

CDS  
1. .1428  
/gene="At4g11920"  
/codon\_start=1

/evidence=experimental  
/product="putative Swi1 protein"  
/protein\_id="AAM14174.1"  
/db\_xref="GI:20259635"  
/translation="MEEDSTPTKKKSQDLNPPSMRPTVLSRLNRLIDSNRYH  
SRSKPIYSRPFPSRSGSNFALPDLASSPNKQKEDGAGSYASLKLALPGPVNPE  
KSDVNVGFPSPGNIRFKETQSRNLNPPFSDVVGVSFSPFKSLRPFYKVL  
DAPLQDDPYLNLVMSQNLVAGLVGNLWNAQNSKVKLCLDGLVDETVCSGVA  
LRTGLHATISSTGQIWDVLRCKNIRMEHRLRVGALWSSVSSGRDSILOR  
DRTQDHYSLKLGKSEICGLKNSRDELASGNDKLFVNHQHSIQPLRFCEHA  
AAVKAATWSPHFGLLAGGGGADRCIRFWNTTTHLNCVDVINSQVNLVMSKNVE  
LYSTHGSQNIIVKQYPMKSLATLTHGSHRYLVLYAVSPDQIIVTGAGDETLRFVN  
VPPSPKSQSEIEGALSFGRTIIR"

3'UTR  
1429. .1459  
/gene="At4g11920"  
BASE COUNT 393 a 287 C 366 g 413 t  
ORIGIN

Query Match 35.3%; Score 708.4; DB 8; Length 1459;  
Best Local Similarity 71.7%; Pred. No. 2.6e-155;  
Matches 957; Conservative 0; Mismatches 371; Indels 6; Gaps 2;

Qy	276	TAAGCCGATGATCAACAGCAACATACACCTCCCTCTCGAACAACTACTCCGATA	335
Db	101	TCAACGGTTAATCGATTGGAATCATATCACTCTCTCTCGAAACCTATTACTCAGATA	160
Qy	336	GGTTCAATCCGAGTAGATCTGCTCGAAATTGCTTTGTTTGATATCAATCTCCGACAG	395
Db	161	GGTTTATCCAAAGTAGATCTGGTTCCAAATTCGCGCTTTTCGATTGGCAATCTTCGCG	220
Qy	396	AAGGACCGGATGATGTTCCAGCGCTTATACGACTCTTCTGAGAACGGCGTTGTTTGAC	455
Db	221	CGAATAAGAAGATGG---GAAGAAGATGGGCTGGTCTTATGCGAGTCTTTTGAAGA	277
Qy	456	CGGATGTTGCGGTCGGGTTACGCCGGAATAAACCAGACTCGCGGTCGATGCAATGCCGA	515
Db	278	CGGCGCTTTTGGTTCGGTCAACCGCGAGAAAAGTGAATGTTGTTAATGGGTTTTCGCGT	337





Db 2774 CGCGTCGATGACATTCGCGATAGGATATTTTAGGTATAGACGAGACAGT 2833  
 QY 555 CAATGACATCGCTTTTCGCGTTTATGAGATGATTTTTCCTGGTGTAAATCATAGTC 614  
 Db 2834 CAATGACATCGCTTTTCGCGTTTATGAGATGATTTTTCCTGGTGTAAATCATAGTC 2893  
 QY 615 CGGTTAAGGCTCCTAGGAAGGTTCTCGATCGCCTTATA----- 653  
 Db 2894 CGGTTAAGGCTCCTAGGAAGGTTCTCGATCGCCTTATAGGTGAATTCATAAAAAAC 2953  
 QY 654 ----- 653  
 Db 2954 TTATCTTTTATTTCTTTGATCTTTGTTGAATGTTGTTAGGTTTGTGATTT 3013  
 QY 654 -----AGGTTTGGATGACCTGCTTTGCAAGATGATTTTATCTGAACTCTGTAGATT 707  
 Db 3014 TGATATAGTTTTCGATGACCTGCTTTGCAAGATGATTTTATCTGAACTCTGTAGATT 3073  
 QY 708 GGTCTTCACAAATGTTGGCTGTGGTTGGTGAATGTTGTTGCTATCTCTGGAATGCTT 767  
 Db 3074 GGTCTTCACAAATGTTGGCTGTGGTTGGTGAATGTTGTTGCTATCTCTGGAATGCTT 3133  
 QY 768 GTAGCAGCAAGTAACTAAATATATGTGAT 796  
 Db 3134 GTAGCAGCAAGTGTGAGTTTATATCATAT 3162

## RESULT 8

BT002165  
 LOCUS 1511 bp mRNA linear PLN 19-NOV-2002  
 DEFINITION Arabidopsis thaliana cell cycle switch protein (At5g13840) mRNA, complete cds.  
 ACCESSION BT002165  
 VERSION BT002165.1 GI:25084104  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1511)  
 Nguyen,M., Karlin-Neumann,J., Southwick,A., Tripp,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banth,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinzaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submision  
 Submitted (19-NOV-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W., Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Torumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.  
 Location/Qualifiers  
 1. .1511  
 /organism="Arabidopsis thaliana"

## FEATURES

source

/mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="U24960"  
 /note="This clone is in pENTR/SD-dTopo. This is a cloned PCR product using RIKEN clone RAFL09-66-B04 (AY099581), as a template  
 ecotype: Columbia"  
 1. .1511  
 /gene="At5g13840"  
 1. .1446  
 /gene="At5g13840"  
 /codon\_start=1  
 /product="cell cycle switch protein"  
 /protein\_id="AA072176.1"  
 /db\_xref="GI:25084105"  
 /translation="MASPQSTKTLNLPAGMQTSLRLTFSSFRGISLSLSPKSKST QGSAFSPFMSFPCNMLRFKTDNSNSPSPFSSILGNDNHSDDSPPPKPRKVPK TPVKVLDAPSLDDPFLVNVDSQNVLAIGLGTCTVYLTASNSKVTKLCDLGPNDV CSQMTREGSYISIGTSHGQVQWDDGTQCKRVRTMGHQTGTGLVLAWSRILSSQSRD RNILQHDIRVQSDPVSFKLVHKSSEVCGKSHDDRELASGNDNOLLVNNHSQOPIIL KLTEHTAAVKAITWSPHSSLLASGCGGTADRCIRFWNTTNGQLNSIDTGSQVQCNLAW SKVNEIVSTHGYSONQIMLWKYPSKVAITGHSRVLVYLATSPDGTITVTGADGE TLRPNVFPSPVQTPVKTGSLSLGRTQIR"  
 BASE COUNT 411 a 342 c 369 g 389 t  
 ORIGIN  
 Query Match 23.5%; Score 472.4; DB 8; Length 1511;  
 Best Local Similarity 66.1%; Pred. No. 5.7e-100;  
 Matches 680; Conservative 0; Mismatches 349; Indels 0; Gaps 0;  
 QY 582 ATGATGATTTTGTCTCGGTGTTAATCATAGTCGGGTAAAGGCTCCTAGGAGGTTCTTC 541  
 Db 419 ATGATAATGTCACCTCCAGTGACTGCTCCGGCTCCGAAGCCCCCTCCAGAGTTCCTA 478  
 QY 642 GATCCCTTATAAGGTTTGGATGACACCTGCTTTTGAAGATGATTTTATCTGAATCTGG 701  
 Db 479 AAACACCTCATAGGCTTTCGATGCTCCCTCTTTTACAAGATGACTTCTACTTGAATGTTG 538  
 QY 702 TAGATGCTCTTACACAAATGCTTGGCTGTGGTTTGGTGAAGTCTGCTATCTCTGGA 761  
 Db 539 TGGACTGAGTTTACAGAAATGTTCTTGGCGTTGGGCTTGTGTTGTTTATCTTTGGA 598  
 QY 762 ATGCTTTAGCAGCAAGGTAACATAATATGTAATTTGGGGTTGATGATTTGTTGTTT 821  
 Db 599 CGGCTTCCAATAGCAAAAGTGACTAAGTATATGACCTAGGCGCTAATGACAGTGTGCT 658  
 QY 822 CTGTTGCTGGGCTCAACGCTGCTCATCTTCTGCTGCTGGAACCTAACATGCTTAAGTTC 881  
 Db 659 CGGTTCACTGACACAGGAGGTTCTATATATCTTATGTTCAAGTCTAGGTCAAGTTTC 718  
 QY 882 AGATTTGGGATGACGCAAGATGCAAGAAGATGAATCAATGGAGGGCCATCGTTACGTC 941  
 Db 719 AGGTTTGGGACGGAAACACAGTGCAGAGAGAGTCCGAACCATGGAGGTCAATCAACAAGAA 778  
 QY 942 TCGGGGCTTGGGCTCGAGGTTTCATCTTTTCTCTCTGCTGACGGGATAGATATTTT 1001  
 Db 779 CTGGTGTCTTGGCATGGAAGTCTAAGATCTTATCATCAGGAGCGAGACAGAAATCTCC 838  
 QY 1002 ATCAACGAGATATACGACACAAAGAGATTTTGTGTAGTAACTGTTCAGGACACAAATCAG 1061  
 Db 839 TTCAAGCATGATATCCGAGTCCAGAGCGATTTTGTTCAGCAAACTCGTGGGCCACAAATCTG 898  
 QY 1062 AGGTTTGTGACGAGAGTGTGTATATGATAACCGTGTGAGTGTGTCATCTGAGGAAATGACA 1121  
 Db 899 AAGTCTGGGCTCTGAATGTTCTCATGATGATAGAGAGCTTCGGTCTGGAGCAATGATA 958  
 QY 1122 ACAATTTGTTTGTGGAATCAACACTCAACCCAGCTGTCTCTCAAGTACTGTGAGACACA 1181  
 Db 959 ATCAGTTTATTTGGTATGGAACAAATCATTTACAGCAACCTATTCTGAAGCTGACTGAGCATA 1018  
 QY 1182 CAGCAGCTGTATAAGCTATTGTCATGCTCTCTCATCTTCTCATGGAATCTTCTTGCATCTGGAG 1241



```

1019 CAGCAGCGGTTAAGGCAATTACATGGTCTCCTCATCAGCAGCGCTCCTGCTTCGGGAG 1078
1242 GAGGAACCTCAGATAGATGTTCTGTTTTCGTAATACACCAACCACTACACCTTAGCT 1301
1079 GTGGAAACCGCAGACAGATGATGATTTCTGGAACACGAAACGGAATCAGTTAAACA 1138
1302 GTATGGACACTGGAGTGGAGTTGCAATCTGTGTCCTGATCCAAATGTCACCAACTAG 1361
1139 GCATCGACACTGGAGCGCAGGCTGCAATCTGGCGTGAGCAAGAAATGTTAACGAGATAG 1198
1362 TAAGCACACATGGGTACTCCAGACACAGATATGTTTGGAGATACCCCACTATGCA 1421
1199 TGAGTACTCATGATGATCTCTCAAAACCAATCATGCTCTGGAAGTACCCTCATGTC 1258
1422 AGCTGGGCACTCTTACCGGCCATATCTATAGGTTTCTCTATCTTGGCCATCTCTCCAGATG 1481
1259 AGTTTGAACACTTACCGGCCACACGATGAGAGTACTTACTTGGCTACGTCACCTGATG 1318
1482 GACAGACTATTGTAAGTGGAGTGGAGTGAACCGCTTAGTCTGGAATGTTTCCCTT 1541
1319 GCCAGACTATAGTGGACCGGAGCAGAGATGAGACCTCGGGTTTGGAACTGCTTTCCTT 1378
1542 CCCCTAAATCAGAGATACTGAAAGTGAATCGAGCACTTATCTCTTGGAAAGAACTACTA 1601
1379 CAGTAAATGACAGACACAGTGAAGACACAGGCTCTGCTGTCATGCGGAGGACACAGA 1438
1602 TCAGGTGAT 1610
1439 TCCGATAAT 1447

```

## RESULT 9

AY099581

LOCUS

DEFINITION

Arabidopsis thaliana cell cycle switch protein (At5g13840) mRNA, complete cds.

ACCESSION

AY099581.1

VERSION

GI:20466230

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

AUTHORS

Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.

Direct Submission

Submitted (24-APR-2002)

DNA Sequencing and Technology Center,

Stanford University, 855 California Avenue, Palo Alto, CA 94304,

USA

e-mail for correspondence: arab@sequence.stanford.edu

COMMENT

The Salk, Stanford, PDEC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs (RAPL cDNA: 'RIKEN

collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PDEC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs (RAPL cDNA: 'RIKEN

Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,

Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K.,

Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P.,

Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed

equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.

(SSP/Stanford) contributed equally to this work as PIs.

## FEATURES

## source

Location/Qualifiers

1..1891

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAPL09-66-B04 (R24960)"

/note="This clone is in pBluescript

ecotype: Columbia"

1..1891

/gene="At5g13840"

286..1731

/gene="At5g13840"

/cdon\_start=1

/product="cell cycle switch protein"

/protein\_id="AA20433.1"

/db\_xref="GI:20466231"

/translation="MASPQSTKTLNLPAGMNTSLRLTFSSFRGSSLSPPSKST

CSDRFPCRSRLHAFDQKPTVPVKGNEAYSRLLKSELFGSDPSPKPKVPK

GOGSASSPMSPTNMLRPKTDNRSSSPSPFPSPILNDNGHSSDSSPPKPKVPK

THKVLDAFSLQDDPYLVNVDMSNVLAAGLGTCTVLTASNTKLCDLGPNDSV

CSVOTREGSYISIGTSHQGVQVWDGTQCKVRVTGGHQTGTGLVANSRIILSSSRD

RNLEQHDIVQSDPFVSKLVGHKSEVCGCLKWSHDDRELASGNDNQLLWNNSIQPIL

KLTEHTAAVKAITWSPHOSLLASGGGTADRCIRFNFTNTGNDLNSIDTGSQVCLAW

SKVNEIVSTHGYSQNIIMLKYPSMKVATLTTHGSMRYLYLATSPDGTITVAGDDE

TLRFNVFSPVKMTFVKDTGLWSLGRITQIR"

BASE COUNT 537 a 427 c 422 g 505 t

## ORIGIN

Query Match 23.5%; Score 472.4; DB 8; Length 1891;

Best Local Similarity 66.1%; Pred. No. 5.9e-100;

Matches 680; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

Qy 582 ATGATGATTTTGTTCCTGGTGTAAATCATAGTCCGGTTAAGGCTCTCTAGGAAGTTCCTC 641

Db 704 ATGATAATGTCTACTCAGTGACTCGTCCGCTCGAAGCCCTCGCAAGTTCCTA 763

Qy 642 GATCGCTTATAAGTTTGGATGCACTGCTTTGCAAGATGATTTTATCTGAATCTGG 701

Db 764 AACACCTCATAGTCTTGGATGCTCTCTTTTACAAGATGACTCTTACTTGAATGTTG 823

Qy 702 TAGATTGGTCTTACACAAATGTTGGCTGTTGGTGTGCTTCTATCTCTGGA 761

Db 824 TGGACTGGAGTTCACAAATGTTCTTCCGCTGGGCTTGGTACTTGTGTTTACTTTGGA 883

Qy 762 ATGCTTCTAGCAGCAAGTAACTAAATATGTTATGTTGGGGTTGATGATGTTTGT 821

Db 884 CGGCTTCCAATAGCAAGTGAAGTAAAGTTATGTGACCTAGGCGCTTAATGACAGTGTGCT 943

Qy 822 CTGTTGGTGGCTCAACGTGTTACTCATCTTCTGTTGGTGAACCTAACAAATGTTAAAGTTC 881

Db 944 CGGTTCAAGTGACACCGGGAAGGTTTCATATATCTATTTGGTCAAGTCAAGTTCAGGTTTC 1003

Qy 882 AGATTGGGATGACGCAAGATGCAAGAGATGAAGATCAATGGAGGGGCTATCGGTTACGTG 941

Db 1004 AGTTTGGGACGGAACACAGTGAAGAGTCCGAACCATGGGAGGTTCATCAACAAGAA 1063

Qy 942 TCGGGGCTTGGCTGGAGTTCACTCTTTTCTTCTTCTTGGTGGACGGGATAGATATTT 1001

Db 1064 CTGGTGTCTTGGCATGGAACTCAAGGATCTTATCATCAGGGAGCAGACAGAAATCC 1123

Qy 1002 ATCAACGAGATATACGACACAGAAATTTGTTAGTAAACTGTCAAGGACACAAATCAG 1061

Db 1124 TTCAGTATGATTCGAGTCCAGAGCGATTTTGTGACAACTCGTGGGCGACAAATCTG 1183

Qy 1062 AGGTTTGGGACTGAAGTGGTCATATGATACCTGATGTTGGCATCTGGAGGAAATGACA 1121

Db 1184 AAGTCTGGGCTGCAAAATGGTCTCATGATGATGAGAGCTTGGCTCTGGAGGCAATGATA 1243

Qy 1122 ACAATTTGTTTGGTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTGAGACACA 1181

Db 1244 ATCAGTTATTTGGTATGGAACTATTCACAGCAACCTATTTCTGAGCTGACTGAGCATA 1303



through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 96 Row: k Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis. Similarity but not identity to protein.

FEATURES source

```

1. 3033
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cbone="MGC:53225 IMAGE:55"
/tissue_type="Embryo, stage"
/clone_lib="NICHHD_XGC_Emb4"
/lab_host="DH10B"
/note=/vector: pCMV-SPORT6
161. 634

```

CD5

BASE COUNT	842 a	615 c	659 g	917 t
ORIGIN				

```
Query Match      20.2%; Score 405.2; DB 5; Length 3033;
Best Local Similarity 66.2%; Pred. No. 3.3e-84;
Matches 616: Conservative 0; Mismatches 308; Indels 6; Gaps 2
```

613	Qy	TCGGTTAAGGCTCCTTAGGAAGGTTTCCTCGATCCCTTATAGGTTTTGGATGACACCTGC	678
736	Db	TCCACGAAGGCTACAAGGAAGATTTCCGAAGATCCCTTTAAAGTTCTGGATGACACCTGA	795
673	Qy	TTTCCAGAATGATTTTTTATCTGAAATCTGGTAGATTGGTCTTTACACAATGTGTGGCTGT	732
796	Db	GTTTACAAGATGACTTTTACCTCAACCTAGTAGGATTTGGTCATCCCTCAATGTTTTAAGTGT	855
733	Qy	TGGTTTGGGTAACTGTGTCTATCTCTGGAAATGCTTTGTAGCAGCAAGGTAACATAATTATG	792
856	Db	GGGCTTGGGAACTTGTGTGTAATTTATGGAGTGCATGTACTAGCCAGAGTGACTAGACTATG	915
793	Qy	TGA---TTTGGGGGTTGATGATCTGTGTTGTTGTTGGTTGGGCTCAACCTGGTACTCA	849
916	Db	TGACCTCTCTGTGAAAGGGGAATCTGTAACTTTCAGTTGGTTGGTCAGAGCGGGGAAACCT	975
850	Qy	TCTTGCTGTTGGAACCTAAACATGCTAAAGTTTCAGATTTGGGATGCAGCAGATGCAGAA	909
976	Db	GGTAGCTGTGGGGACACACAAGGCTTTGTTTCAGATTTGGGATGCATCAGCAGGGAAGAA	1033
910	Qy	GATAAGATCAATGAGAGGCCATTCGGTTACGTGTGGGGCCCTTGGCTGGAGTTCACTCT	969
1036	Db	ACTGTCCACTCTGGAGGGACACACAGCCCTGTGGGACCTTGGCTTGGAAATGCAGACCA	1099
970	Qy	TTTGTCTTCGTTGGAGCGGATAGAATATTTATCAACGAGATATACGCACACAA--GA	1022
1086	Db	GCTTTCCTCTGGAGTAGGGACAGATGATCTTACAAGGGATATTCGCAACCACTGT	1153
1027	Qy	AGATTTTCTTAGTAACTGTGAGCACAAATTCAGAGGTTTGTGSACTGGAAGTGTGCATA	1088
1156	Db	TCAGTCAGAGAGAAGCTGCAAGGCCACAGGCAAGAAGTGTGTGGCCTCAAGTGGTCCAC	1211
1087	Qy	TGATAACCGTAGTTGGCATCTGGAGGAAATGACAACAAATGTTGTTTGGAAATCAACA	1144
1216	Db	AGACCACCAAGCTCCCTGGGCTTCAGGAGGGAATGACAACCAAGCTCTTGGTTGGAAACCATC	1271
1147	Qy	CTCAACCCAGCGTGCCTCAAGTACTGTGTAGCACACAGCAGCTGTTTAAAGCTATTGCAATG	1204
1276	Db	TAGCCTGAGTCTCTGTTCAACAGTATACGTGAACATCTAGCTGCAGTGAAGCAATTTGCCCTG	1333
1207	Qy	GTCTCCTCATCTCATGCACTTCTTGCATCTGGAGGGAACCTGCAGATGATGATTTCG	1266

clone distribution: MGC clone distribution information can be found

Db 1336 GTCCCTCATCAACATGGCTTTTGGCTTCAGGCGGGGCACTGCTGACCGCTGCATCG 1395  
 QY 1267 TTTTGGATACACACAACTCACACCTTAGCTGTATGGACACTGGAGTCAGTTTG 1326  
 Db 1396 CTTCTGGAACACATACAGGCGAGCTTTTACAGTGTATTGATCTGGCTCTCAAGTCG 1455  
 QY 1327 CAATCTTGTCTGTCGTCACAAAATGTCACGAACCTAGTAAGCACACATGGTACTCCAGAA 1386  
 Db 1456 TAACTTGGCTGCTTAAGCATGCAATGAACCTAGTAAAGCACACATGGCTACTCTCAAAA 1515  
 QY 1387 CAGATTATTGTTGAGATACCCACTATGTCAAAGCTGGGCACTTTACGGGCATAC 1446  
 Db 1516 TCAGATTATTAGTATGAATATCCATCGCTTACACAGGTGGCCAAAGTTGACTGGACATTC 1575  
 QY 1447 TTATAGGTTCTCTATCTTCCATCTCTCCAGATGACACAGATATTGTAACCTGGAGCTGG 1506  
 Db 1576 CTACCGTGTCTCTACTCGCTATGTCACAGATGGGAGGCCATTGTTACAGGAGCAGG 1635  
 QY 1507 AGATGAACCGCTAGGTTCTGGAATGTTTT 1536  
 Db 1636 TGATGAACGCTGAGATTCTGGAATGTTTT 1665

## RESULT 11

XLNFRP 2649 bp mRNA linear VRT 08-SEP-1997  
 LOCUS  
 DEFINITION  
 XENOPUS LAEVIS MRNA FOR FIZZY-RELATED PROTEIN.  
 ACCESSION  
 Y14163  
 VERSION  
 Y14163.1 GI:2326942  
 KEYWORDS  
 fizzy-related protein; fzr gene.  
 SOURCE  
 XENOPUS LAEVIS (AFRICAN CLAWED FROG)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Fipidae;  
 Xenopodinae; Xenopus.

## REFERENCE

1 Sigris S.J. and Lehner, C.F.  
 Drosophila fizzy-related down-regulates mitotic cyclins and is  
 required for cell proliferation arrest and entry into endocycles  
 Cell 90 (4), 671-681 (1997)

## JOURNAL

MEDLINE  
 PUBMED  
 97433078  
 9288747

## REFERENCE

2 (bases 1 to 2649)

## AUTHORS

Lehner, C.F.

## TITLE

Direct Submission

## JOURNAL

Submitted (04-JUL-1997) C.F. Lehner, Department of Genetics,  
 University of Bayreuth, 95440 Bayreuth, FRG

## FEATURES

Location/Qualifiers

1..2649

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/tissue\_type="ovary"

1..2649

/gene="fzr"

10..99

/gene="fzr"

100..1581

/gene="fzr"

/codon\_start=1

/product="fizzy-related protein"

/protein\_id="CAA74576.1"

/db\_xref="GI:2326943"

/db\_xref="SPRMBL:O42585"

/translation="MDQYERRLIQINLQNETIPCASMRRTLTPTNSPMSSPSKH

GRFIPSRAGNSINPHRINENEKSPONRKADATADSGKDLAYSALLKNEILGA

GIEKVPQEDERLQSPTEPKSLFTYSLSKRAPDDGNEVSPYSLSPVSNKSKL

LRSPRTKISKIPKVLDAPELODDPYLVNDSIATLVSVGLACTCYVLSACTSQ

VTRCLDSVSGDSVTSVGSERGNLVAVGHKGFVQIWDASAKKLSLTLEHTARVGA

LAWNDQLSSGRDMTLQDRTFPVQSERLQHRQEVCGKLSSTDHQLLASGND

NKLLVNHSSLSVQYTHELAALKAIWSPHQGLIASGGTADRCIRFNTLTQOP

LOCTDGTGQVNCNLAWSKHNELVTHGYSQNLVWKYPSLTQVAKLTGHSRVLYLA

MSPDGEALVTGADETLRFWNVFSKTRSTKESVNLNLFTRIR"

3'UTR 1582..2649  
 BASE COUNT 758 a 541 c 598 g 752 t  
 ORIGIN  
 Query Match 20.1%; Score 403.6; DB 5; Length 2649;  
 Best Local Similarity 66.1%; Pred. No. 7.9e-84;  
 Matches 615; Conservative 0; Mismatches 309; Indels 6; Gaps 2;  
 QY 613 TCCGGTTAAAGCTCCTTAGGAAGGTTCTCGATCGCCTTAAAGGTTTGGATGCACTGC 672  
 DB 588 TCCACGAAGGCTTACAGGAAGATTTCGAAGATCCCTTTAAAGTTCTCGATGCACTGA 647  
 QY 673 TTTCAGAGATGATTTTATCTGAATCTGGTAGATTTGGTCTTCACAGATGTTGGCTGT 732  
 DB 648 GTTACAGAGATGATTTTACCTCAACCTAGTGGATTTGGTCAATCCCTCAATGTTTAAAGTGT 707  
 QY 733 TGGTTTGGGTAACGTGTCTATCTCTGGAATGCTTTGTAGCAGCAAGGTAACATAATATG 792  
 DB 708 GGGCTTGGGAACTTGTGTACTTATGAGTGGCATGACTAGCCAGGTGACTAGCATATG 767  
 QY 793 TGA---TTTGGGGGTTGATGATTTGTTTCTGTTGGTTGGCTCAACGTGTACTCA 849  
 DB 768 TGACCTCTCTGTGAAGGGGATTCTGTAACCTTCAGTTGGTGTGAGCGGGGAAACCT 827  
 QY 850 TCTTGTCTTGGAACTAACCAATGCTAAAGTTTCAGATTGGATGACAGCAAGATGCAAGAA 909  
 DB 828 GGTAGCTGTGGGAAACACAAAGGCTTTGTTCAAGTTTGGATGATCAGCAGGAGAA 887  
 QY 910 GATAAGATCAATGAGAGGCCATCGGTTACGTTGCGGGCCTTGGCCTGGAGTTCACTCT 969  
 DB 888 ACTGTCCACTCTGAGGAGCACACAGGCCGCTGTGGAGCCTTGGCTTGGAAATGAGACCA 947  
 QY 970 TTTGTTCTTCTGGTGGAGGATAGATATTTATCAACAGATATAGCACACAAA---GA 1026  
 DB 948 GCTTTCCTCTGGAGTAGGACAGATGATCTTCAAGGGATATTGACACACCACTGT 1007  
 QY 1027 AGATTTTGTAGTAACTGTGAGACACAAATCAGAGCTTTTGGATGAGTGGTCTATA 1086  
 DB 1008 TCAGTTCAGAGAGAGGCTGCAAGGCCACAGCAAGAAAGTGTGGCCTCAAGTGGTCCAC 1067  
 QY 1087 TGATTAACCGTGAAGTGGCATCTCGAGGAATGACAAATTTGTTTGGTGGATCAACA 1146  
 DB 1068 AGACCCAGCTCTCGGCTTCAGGAGGAAATGACAAAGCTCTTGGTTTGGAAACCATTC 1127  
 QY 1147 CTCACCCAGCTCTCTCAAGTACTGTGAGCACACAGCAGCTGTAAAGCTATTGCAATG 1206  
 DB 1128 TAGCCTGAGTCTCTTCAACAGTATCTGAACTAGCTGAGTGAAGCAATTGCTCG 1187  
 QY 1207 GTCTCTCATCTTATGCACTTCTGATCTGGAGGAGGAAGTGCAGATAGATGATTCG 1266  
 DB 1188 GTCCCTCATCAACATGCGCTTTTGGCTTCAGGCGGGGGCACTGCTGACCCCTGCATCCG 1247  
 QY 1267 TTTTGGAAATCAACACAAACTCACACCTTAGCTGTATGAGACACTGGAAGTCTAGGTTG 1326  
 DB 1248 CTTCTGGAAACACACTGACAGCGCCAGCTTTACAGTGTATTGATCTGCTCAAGTCTG 1307  
 QY 1327 CAATCTTGTCTGTCACAAAATGTCAACGAATAGTAGACACATGGTACTCCAGAA 1386  
 DB 1308 TAACTTGGCTGGTCTAAGCATGCAATGAAATGAACTGGTAAAGCACATGCTACTCTCAAAA 1367  
 QY 1387 CCAGATTATTGTTGGAGATACCCCACTATGTCAAAAGCTGGCGACTCTTACCGGCCATAC 1446  
 DB 1368 TCAGATTATTAGTATGGAATATCCATCGCTACTCAGGTGGCCAAAGTTGACTGGACATTC 1427  
 QY 1447 TTATAGGTTCTCTATCTTGGCACTCTCCAGATGGAACAGACTATTGTAACCTGGAGCTGG 1506  
 DB 1428 CTACCGTGTCTCTACCTGGCTATGTCAACAGATGGGAGGCCATTGTTACAGGAGCAGG 1487  
 QY 1507 AGATGAACCGCTAGGTTCTGGAATGTTTT 1536  
 DB 1488 TGATGAACCGCTGAGATTCTGGAATGTTTT 1517

```

RESULT 12
BC048038      2089 bp      mRNA      linear      VRT 05-MAR-2003
LOCUS         Danio rerio, clone MGC:55450 IMAGE:2639752, mRNA, complete cds.
DEFINITION    BC048038
ACCESSION     BC048038.1 GI:28856199
VERSION       MGC
KEYWORDS      Danio rerio (zebrafish)
SOURCE        Danio rerio
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
REFERENCE     1 (bases 1 to 2089)
AUTHORS       Strausberg R.
TITLE         Direct Submission
JOURNAL       Submitted (04-MAR-2003) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
              cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              contact: amadan@systemsbio.org
              Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
              Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 101 Row: h Column: 12.
Location/Qualifiers
1. 2089
   /organism="Danio rerio"
   /mol_type="mRNA"
   /strain="AB"
   /db_xref="taxon:7955"
   /clone="MGC:55450 IMAGE:2639752"
   /tissue type="Whole body, adult, (one male and one female,
   including unfertilized eggs)"
   /clone_lib="Sugano kawakami zebrafish DRa"
   /lab_host="DH10B"
   /note="Vector: pME186-FL3"
   /codon_start=1
   /product="Unknown (protein for MGC:55450)"
   /protein_id="AAH48038.1"
   /db_xref="GI:28856200"
   /translation="WDQYERLLRQINVDNTSPMKADVLCHLTPTPESPLSPS
   RKGDRFIPRAGANWSINFRINENKSPSNKTKDTSKADSLAYSAALLKREL
   LAGLEKVDLPOTEDRRLOQPTPERRSLFYSLSAKRSTPDNSVSPISLSPSSSQ
   KLRSPRKPRIKIPKVLDAPELODDFYLNLDWSNLVSLVGLGTCVILWSACT
   SQVRLCDLSVEGDSVTSVSGRNLVAVGTHGFPQIWDATAGKKLFALEGHARV
   GALANADLVSSGRDRMLQDRTPPLOSERLQGRHROVCGKLKMSDHLIASGG
   NDKLLVNHSSVLPMQYTERHLAAVKALMSPHOGLLASGGGTADRCIRFWNTLTA
   OPLQCIDTGSQVCLNLSKTNELVSTHGYSONQLVWKYPSLTVAKLAGHSRVLY
   LAMPDGRALVTGAGDETLRFNWNFSKTRSTKESVSVINLFTIR"
BASE COUNT   599 a 495 c 467 g 528 t
ORIGIN
Query Match      17.8%; Score 357; DB 5; Length 2089;
Best Local Similarity 63.4%; Pred. No. 6.7e-73;
Matches 581; Conservative 0; Mismatches 330; Indels 6; Gaps 2;

QY 647 CTTTAAAGTTTGGATGACCTCTTTGCAAGATGATTTTATCTGAATCTCGTAGAT 706
DB 755 CGGTAAAGTTCCTTGATGCACAGAGCTACAGGACGACTTCTACCTCAACTTAGTGAC.814
QY 707 TGGTCTTCACAAATGTGTGGCTGTGTTGGGTAACATGTGTATCTCTGGAATGCT 766

```

```

DB 815 TGGTCTCTTTAAATGTTCTCAGTGTGGGTCTGGAAACATGTGTTTACTTATGAGTGCC 874
QY 767 TGTAGCAGCAGGTAACTAAATTATGTGA---TTTGGGGTTGATGATGTTGTTGTCT 823
DB 875 TGCACCATCTCAAGTGAACGGCTATGTGACCTATCAGTGGAGGGCGATTCGTCACTCT 934
QY 824 GTTGGTTGGGCTCAACGCTGTPACTCATCTTCTCTTTGAACTAAACAATGGTAAAGTTGAG 883
DB 935 GTGGCTGCTCAGAGAGAGGGGAAATCTTGTCTGTGGGAAACACACAAAGGCTTTGTACAA 994
QY 884 ATTTGGGATCAGCAAGATGCAAGAGATAGATCAATGGAGGGCCATCGGTTCAGGTGTC 943
DB 995 ATTTGGGATGCCCGGCTGGCAAAAACCTGTTGCTTGGGGGACACACTGCAAGATT 1054
QY 944 GGGGCTTGGGCTGGAGTTTCATCTCTTTTGTCTTCTGTGGACGGGTAAGAATATTTAT 1003
DB 1055 GGTGCATTAGCATGGAATGCAGATCAGTTGTCAACAGGCAGTCTGTGACAGATGATCCTC 1114
QY 1004 CAACAGATATACGCA---CACAAGAGATTTTGTAGTAACTGTCAAGGACACAAATCA 1060
DB 1115 CAGAGAGATATCAGGACTCCCTCTCTGCAATCAGAGCGCAGGCTGCGGGCCACAGACAG 1174
QY 1061 GAGGTTTCTGAGTCAAGTGTTCATATATCAATCAACCGTGAAGTTGGCATCTGGAGGAAATGAC 1120
DB 1175 GAGGTTCTGGGACTCAAGTGGAGCAGTCAACCATCAGCTGCTCGCTCTGGGGGAAATGAC 1234
QY 1121 AACAAATGTTTGTGGAAATCAACACTCAACCCAGCTGTCTCAAGTACTGTGAGCAC 1180
DB 1235 AATAAGCTGTTAGTGTGGAATCACTCAAGTGTGTACCAATGCAGCAGTCACTGAAAC 1294
QY 1181 ACAGCAGCTGTTPAAAGCTATTGATGTTCTCTCATCTTCTCATGGAATCTTTCGATCTGGA 1240
DB 1295 TTGGCAGCAGTGAAGGCCATCGCTGTGTTCTCTCATCAACAGGGCTACTGGCGTCTGGA 1354
QY 1241 CGAGAACTGCAGATAGATGATTCGTTTGTGGAAATACAACAACAACTCACACCTTAGC 1300
DB 1355 GGGGGAAACGGCGGACCGCTGTATTGCTTTTGGAACTCTCACAGCAGCAGCAGTGCAG 1414
QY 1301 TGTATGGACACTGGAGTCAAGTGTTCGTTCTGTTCCAAATGTCAACGAATA 1360
DB 1415 TGCATCGACACCGGCTCTCAAGTCTGCAATCTGGCTGGTCCAAACACACCAATGACTG 1474
QY 1361 GTAAGCACACATGGGTACTCCAGAACAGATATTGTTGGAGATACCCCACTATGTCA 1420
DB 1475 GTCAGCACACACGGTTACTCTCAGAATCAGATCTCTGGTGTGGAATATCCCTCACTACT 1534
QY 1421 AAGCTGGGCACTTTACCGGCCATATATAGGTTCTCTATCTTGGCCATCTCTCCAGAT 1480
DB 1535 CAAGTAGCCAAACTGACCGGCCACTCTTATAGGGTGCTTATCTGGCAATGTCTCCAGAT 1594
QY 1481 GGACAGACTATTGTAACTGGAGCTGGAGATGAACCTTAGGTTCTGGAATGTTTTCCT 1540
DB 1595 GGAGAGCGCATAGTACCGGGAGCAGCGATGAATGAATCTCGCGCTTTTGGAAATGTTATTAGT 1654
QY 1541 TCCCTTAATCAGAA 1557
DB 1655 AAAACACGATCAACAAA 1671

```

```

RESULT 13
BC006616      1861 bp      mRNA      linear      ROD 16-APR-2003
LOCUS         Mus musculus frizzy-related protein 1 (Drosophila), mRNA (cdna
DEFINITION    clone MGC:11723 IMAGE:3967009), complete cds.
ACCESSION     BC006616
VERSION       BC006616.1 GI:13879283
KEYWORDS      MGC.
SOURCE        Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 1861)

```

AUTHORS	Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schenken, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Green, E.D., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinska, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	2238257		
REFERENCE	12477932		
AUTHORS	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (27-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisse, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAC Plate: 17 Row: 1 Column: 12 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. Location/Qualifiers 1..1861 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:11723 IMAGE:3967009" /tissue_type="Mammary tumor. Metallothionien-TGP alpha model. 10 month old virgin mouse. Taken by biopsy." /clone_lib="NCI CGAP Mam1" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1..1861 /gene="Pzrl" /note="synonym: Pzr" /db_xref="LocusID:56371" /db_xref="MGI:1926790" 224..1705 /codon_start=1 /product="Pzrl protein" /protein_id="AAH06616.1" /db_xref="GI:13879284"		
source	1..1861		
gene	Pzrl		
CDS	1..1861		

BASE COUNT	409 a	560 c	562 g	330 t
ORIGIN	Query Match 17.6%; Score 352.8; DB 10; Length 1861; Best Local Similarity 62.9%; Pred. No. 6.4e-72; Matches 581; Conservative 0; Mismatches 337; Indels 6; Gaps 2; QY 620 AAGCTCTAGGAGGTTCTCGATCGCCCTATAAGGTTTGGATGACACCTGCTTGCAG 679 DB 719 AAGCCACACGCAAGATCTCTAAGATTCCCTTCAAGGTCTGGACGGCCAGAGCTTCAG 778 QY 680 GATGATTTTATCTGAATCTGGTAGATTGGTCTTTCACACAATGTGTGGCTGTGGTTTG 739 DB 779 GACGACTTCTACCTCAACTTGGTGGATCTGCTCTCTCTCAACGTCTCAGCGTGGGCTG 838 QY 740 GGTAACTGTGTCTATCTCTGGAATCTGTAGAGCAAGTAATAATTATGTGATTG 799 DB 839 GGCACCTCGGTGATCTGTGAGTGCATGCACACGACGAGTGACCCGGCTCTGTGACCTC 898 QY 800 GGGGTGTA---TGATTGTGTGTGTTGTTGTTGGTGGGCTCAACGTGTGTCTCATCTTGTCT 856 DB 899 TCTGTAGAGGGGACTCAGTGACTCTCGTTGGTGGTCTGAGCGGGGAACTTGGTGCAG 958 QY 857 GTTGAACCTAACTAGTAAAGTTTCAAGTTTGGATGAGCAAGATGCAAGATGAAGA 916 DB 959 GTAGTACACACAAGGGCTTCTGTCAGATCTGGGACGCTGCTGTGGGAAAGAGCTGTCC 1018 QY 917 TCAATGGAGGGCCATCGGTACGTTGCGGGCCCTTGGCTGGAGTTCATCTCTTTTGTCT 976 DB 1019 ATGCTGGAGGGCCACACAGCAGAGTGGGGCGCTGGCTGGATGCTGACCACTGTGCA 1078 QY 977 TCTGTGACGGGATAGAATATTATCAACAGAGATATACGACACAA---GAAGATTTT 1033 DB 1079 TCTGTAGCCGTGACCGCATGATCTCAGCGGGATATCCGACACACACCCCTCGCAGTCA 1138 QY 1034 GTTAGTAACCTGTACGAGCACAAATCAGAGTTTGTGGACTGAGTGTGATATGATAAC 1093 DB 1139 GAGCGCGGCTGACAGGGCCACCGGAGGAAGTGTGGCCCTAAAGTGTGCACAGACCAC 1198 QY 1094 CGTGAGTTGGCATCTGGAGGAAATGACAAACAAATTTGTTTGGAAATCAACACTCAACC 1153 DB 1199 CAGCTGCTTGTCTCGGGGGCAATGACAAACAGCTGCTGTGGAACTCTAGTCTA 1258 QY 1154 CAGCTGCTTGTCTGAGGACACAGCAGTGTAAAGCTATTGATGCTGCTCTCT 1213 DB 1259 AGCCCTGTGACAGCATATACGAGGACCTGCGCGGCTGAAAGGCTATTGCTGGTCCCA 1318 QY 1214 CATCTTCATGACTTCTTTCATCTGGAGGAGAACTGCAGATAGATGATTTCTGTTTGG 1273 DB 1319 CACAGCATGACTCTGTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1378 QY 1274 AATACAAACCAAACTCAACCTTAGCTGTATGGACACTGGAAGTGCAGGTTTGCATCTT 1333 DB 1379 AACACTCTGACAGGTGACGCCACTGCGAGTGCATTGACACAGGCTCAACAGTGTGCAACTG 1438 QY 1334 GTCTGTCACAAATGTCACAGAACTAGTAAGCAGACATGGTGTCTCCAGAACAGATTT 1393 DB 1439 GCCTGTCACAGCAGCCCAATGAGCTGTGAGCAGACATGCTACTCAGACAGCAGATC 1498 QY 1394 ATTGTTGGAGATACCCCACTATGTCAAAGCTGGGAGCTCTTACCGGCCATCTTATAGG 1453 DB 1499 CTCGTGTGAGAGTACCCGCTCCCTTACGCGAGGTGGCCAGCTCACTGGCCACTCATATGCT 1558			

QY 1454 GTTCTTATCTTCCATCTCTCCAGATGGACAGACTATTGTAACTGAGCTGAGATGAA 1513  
Db 1559 GTCTCTTACCTGCTGCTCCCTGATGGGAGGAGCCATAGTACCGGAGCTGAGATGAG 1618  
QY 1514 ACCTTAGGTTCTGGATGTTTC 1537  
Db 1619 ACCTGAGGTTCTGGAATGTTTC 1642

RESULT 14  
AF083809 2258 bp mRNA linear ROD 01-SEP-1999  
DEFINITION Mus musculus fizzy-related protein (Fyr) mRNA, complete cds.  
ACCESSION AF083809  
VERSION AF083809.1 GI:5813824  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Jin, D.-Y. and Jeang, K.-T.  
1 (bases 1 to 2258)  
TITLE Characterization of mouse fizzy-related protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2258)  
AUTHORS Jin, D.-Y.  
DIRECT SUBMISSION  
TITLE Submitted (12-AUG-1998) LMM, NIAID/NIH, 9000 Rockville Pike,  
JOURNAL Bethesda, MD 20892-0460, USA  
FEATURES  
Location/Qualifiers  
1..2258  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..2258  
/gene="Fyr"  
143..1624  
/gene="Fyr"  
/function="cell cycle regulator"  
/codon\_start=1  
/product="fizzy-related protein"  
/protein\_id="AA052029.1"  
/db\_xref="GI:5813825"

gene  
CDS

BASE COUNT 463 a 702 c 656 g 437 t  
ORIGIN  
translation="MDQYERLRRLQRIIINQNTVPCVSEMRRLTLPANSPVSPSKH  
GDRFIPRAGANWVFNHINENEKSPQRKADATDNKGKGLAYSLKLLNELGA  
GIERVQPPQEDRLQPTPEHKLFTYLSLSKSSPDNDVSPVSLSPVSKSKL  
LRSRKPRKTSKIPKVLDAPELDQDFYLVNWSLNLVSLVGLGTCVYLWSACTSQ  
VTRLDLSVSGDSVTSVGSERGLVAVGTHKGFQVWDAAGKGLSMLRGHTARVGA  
LAWADQLSSGSRDMILORDIRTPLOSERLOGHREVCGLKWSLTHOLLASGND  
NKLIVNHNSSLPVQQTTEHLAAVKAIAWSPHQHLLASGGGTADRCIRFWNTLTGP  
LOCIDTSGVNCNLAWSKHANLSTHGTFSQNLILVWKFPSLTQVAKLTHSYRLVLA  
MSPDGEAIVTGAGDETLRFMNVFSKTRSKESVSLNLFTRIR"  
Query Match 17.5%; Score 351.2; DB 10; Length 2258;  
Best Local Similarity 62.8%; Fred. No. 1.5e-71;  
Matches 580; Conservative 0; Mismatches 338; Indels 6; Gaps 2;  
QY 620 AAGGCTCTAGGAAGGTTCTCGATCGCTTATAGGTTTGGATGACCTGCTTGCAA 679  
Db 638 AAGCCACAGCAGCAAGATCTTAAGATTCCTTCAGGTGCTGGACGCCAGAGCTTCAG 697  
QY 680 GATGATTTTATCTGAATCTGTAGATGGTCTTACACATGTTGCTGTTGTTG 739  
Db 698 GACGACTTCTACCTCAACTTGTGACTGGTCTCCCTCAACGTGCTCAGCGTGCGGCTG 757  
QY 740 GGTAAGTGTCTATCTCTGGAATGCTGTAGCAGAGGTAACATAATTATGATTG 799  
Db 758 GGCACCTGCTGCTACTCTGGAGTGCATGCACACGAGGTGACCGGCTCTGTGACCTC 817  
QY 800 GGGGTGA---TGATTTGTTTGTCTGTTGGTGGCTCAACGTTGACTCACTTGTCT 856

Db 818 TCTGTAGAGGGGACTCAGTGACTTCGTTGGCTGGTCTGAGCGGGGAACTTGGTCGCA 877  
QY 857 GTTGAACCTAACATAGTAAAGTTTCAGATTTGGATGTCGACAGATGTCAGAGATGAAGA 916  
Db 878 GTAGGTACACAAAGGGCTTCGTGCAGATCTGGACGCTGCTGCTGGAGAGAGCTGTCC 937  
QY 917 TCAATGGAGGGCCATCGTTACGTGTCGGGCGCTTGGCTGGAGTTTCATCTCTTTTGTCT 976  
Db 938 ATGCTGGAGGGCCACACAGACAGAGTGGGGCGCTGGCTGGATGCTGACCACTGTGCA 997  
QY 977 TCTGGTGAACGGGATAGATATTATCAACGAGATATACGACACAA---GAAGATTTT 1033  
Db 998 TCTGTGATCCGCTGACCGCATGATCTTACGCGGATATCCGACACACCCCTCGAGTCA 1057  
QY 1034 GTTAGTAACTGTCCAGGACACAAATCAGAGGTTTGTGACTTGAAGTGGTCTATATGATAAC 1093  
Db 1058 GAGCGCGGCTGCGAGGCCACCGCAGGAGTGTGTGGCTTAAAGTGGTCCACAGACCAC 1117  
QY 1094 CGTAGTTGGCATCTGGAGGAAATGACAAATAATTTTGTGGATCAACACTCAACC 1153  
Db 1118 CAGCTGCTTGGCTCGGGGGCAATGACAAAGCTGCTCGTGTGGAAACCACTCTAGTCTA 1177  
QY 1154 CAGCTGCTTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCAATGCTCTCT 1213  
Db 1178 AGCCCTGTGACGAGTATACGAGCACCTGSCAGCGCTGAGGCTATTGCTGCTCCCA 1237  
QY 1214 CATCTTCAAGACTTCTTGCATCTGGAGGAGAACTGACAGATAGATGATTGCTTTTGG 1273  
Db 1238 CACCAGCATGAGCTGCTGGCATCTGGTGGCAAGCTGACCGCTGCATCCGATTTCTGG 1297  
QY 1274 AATCAACCCACAACTCAGCTTATGATGACACTGGAAGTTCAGTTTGCATCTT 1333  
Db 1298 AACACTGTGACAGGTGACGCCACTGCACTGCATTTGACACAGGCTCACAAGTGTGCACTG 1357  
QY 1334 GTCTGCTCCAAAATGTCAAGCACTAGTAAGCACACATGGGTACTCCAGAACAGAT 1393  
Db 1358 GCCTGTCCAAAGCACGCCAATGAGCTGGTGGACACACATGGCTACTCAGAACACAGATC 1417  
QY 1394 ATTGTTGGAGTACCCACTATGTCAAGCTGGCGCTTACCGGCCATCTATATAGG 1453  
Db 1418 CTCGTGTGGAAGTACCCCTTCCTTACGAGGTGGCAAGCTCCTGCGCATCTCATATG 1477  
QY 1454 GTTCTCTATCTTCCCATCTCTCCAGATGACAGACTATTGTAATGAGCTGGAGATGAA 1513  
Db 1478 GTCTCTTACCTGGCCATGCTCCCTGATGGGAGGCCATAGTACCGGAGCTGGAGATGAG 1537  
QY 1514 AGCTTAGGTTCTGGAATGTTTC 1537  
Db 1538 ACCCTGAGGTTCTGGAATGTTTC 1561

## RESULT 15

AB013462  
LOCUS  
DEFINITION Homo sapiens mRNA for Pzrl, complete cds.  
ACCESSION AB013462  
VERSION AB013462.1 GI:6463678  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
1  
AUTHORS Kotani, S., Oyama, T. and Todokoro, K.  
TITLE Human homologue of Fizzy-related protein  
JOURNAL Published only in Database (1999)  
REFERENCE 2 (bases 1 to 1491)  
AUTHORS Kotani, S., Oyama, T. and Todokoro, K.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-1998) Kazuo Todokoro, The Institute of Physical  
and Chemical Research (RIKEN), Molecular Cell Science Laboratory,  
3-1-1, Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
(E-mail: todokoro@etc.riken.go.jp, Tel: 81-298-36-9075,

AB013462 1491 bp mRNA linear PRI 19-MAR-2002  
Homo sapiens mRNA for Pzrl, complete cds.  
AB013462  
AB013462.1 GI:6463678  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Kotani, S., Oyama, T. and Todokoro, K.  
Human homologue of Fizzy-related protein  
Published only in Database (1999)  
2 (bases 1 to 1491)  
Kotani, S., Oyama, T. and Todokoro, K.  
Direct Submission  
Submitted (29-APR-1998) Kazuo Todokoro, The Institute of Physical  
and Chemical Research (RIKEN), Molecular Cell Science Laboratory,  
3-1-1, Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
(E-mail: todokoro@etc.riken.go.jp, Tel: 81-298-36-9075,

Db 1123 GCGGCACAGCTGACCGCTGTATCGCTTCTGGAACACGCTGACAGGACCACTGCGAG 1182  
 Qy 1301 TGTATGGACACTGGAGTCAAGCTTTTGGCAATCTTGTCTGTCCTCAAAATGTCACGAACTA 1360  
 Db 1183 TGTATCGACACCGGCTCCCAAGTGTGCAATCTGGCTGTGTCAGCAAGCAAGCAAGCTG 1242  
 Qy 1361 GTAAGCACACATGGGTACTCCAGAACCAAGATTATTTGTTGGAGATACCCCACTATGTCA 1420  
 Db 1243 GTGAGCACGACGCTTACTCAGAGAACCAAGATCTTGTCTGGAAGTACCCCTCCCTGACC 1302  
 Qy 1421 AGCTGGGACTTTTACCGGCCATCTTATAGGTTCTCTATCTTGGCAATCTTCCAGAT 1480  
 Db 1303 CAGGTGGCCCAAGCTGACCGGGCACTCTACCGCGTGTCTTACTTGGCAATGTCCTGAT 1362  
 Qy 1481 GGCACAGCTATTGTAACCTGGAGCTGGAGATGAACGCTTGTAGTTCTTGGCAATGTTT 1536  
 Db 1363 GGGGAGGCCATCGTCACTGGTCTGAGAGACGAGACCTGAGGTCTTGGAAAGCTT 1418

Search completed: January 23, 2004, 19:53:10  
 Job time : 7472 secs

FEATURES  
 source  
 gene  
 CDS

Location/Qualifiers  
 1..1491  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1..1491  
 /gene="fzrl"  
 1..1491  
 /note="Fizzy-related protein"  
 /codon\_start=1  
 /product="Fzrl"  
 /protein\_id="BAA86954.1"  
 /db\_xref="GI:6463679"

BASE COUNT 312 a 479 c 452 g 248 t

ORIGIN

Query Match 17.0%; Score 340.8; DB 9; Length 1491;  
 Best Local Similarity 62.8%; Pred. No. 4.2e-69;  
 Matches 563; Conservative 0; Mismatches 327; Indels 6; Gaps 2;

Qy 647 CTTTATAAGGTTTGGATGACCTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGAT 706  
 Db 523 CCCTTCAAGGTGCTGGACGCGCCGAGCTGACGAGCACTTCTACCTCAATCTGGTGGAC 582  
 Qy 707 TGGTCTTACACAATGTGTGCTGTTGGTTGGTAACTGCTATCTCTGGAATGCT 766  
 Db 583 TGGTCTGCTCAATGTGCTAGCGTGGGCTAGGCACCTCGTGTATCTCTGGAGTCCC 642  
 Qy 767 TGTAGCAGCAAGTAACTAAATTTATGTGA---TTTGGGGTTGATGATTTGTTGTTCT 823  
 Db 643 TGTACAGCAGGTGACGCGGCTCTGTGACCTCTCAGTGAAGGGCACTCAGTGACCTCC 702  
 Qy 824 GTTGGTTGGGCTCAAGTGGTACTCATCTTGTCTGTTGGTAACTAACTATGTTAAAGTTT 883  
 Db 703 GTGGGCTGCTGAGCGGGGAACTTGGTGGGCTGGGCAACACCAAGGGCTTCGTGGCAG 762  
 Qy 884 ATTTGGGATGACAGATGACAGAGATAAGATCAATGGAGGGCCATCGTTACGTTCTC 943  
 Db 763 ATCTGGGACGACCGCAGGAGAGAGCTGTCATGTTGGAGGGCCACACGCGCAGCTC 822  
 Qy 944 GGGGCTTGGCTGGAGTTCACTCTTTTGTCTTGTGGTGGACGGGATGAAGATTTAT 1003  
 Db 823 GGGGCGCTGGCTGGAATGCTGAGCAGCTGTCTCGGGAGCGCGCATGATCTCTG 882  
 Qy 1004 CACAGGATATACGCA---CACAGAAGATTTTGTAGTAACTGTGAGGACACAAATCA 1060  
 Db 883 CAGAGGACATCCGACCCCGCCACTGCGAGTGGAGCGGCGCTGACAGGGCCACCGGAG 942  
 Qy 1061 GAGGTTTGTGACTGAAGTGGTATGATTAACCGTGAAGTGGCATCTGGAGGAATGAC 1120  
 Db 943 GAGGTGTGGGCTCAAGTGGTCCACAGACCACTCTCTCGCTCGGGGGGCAACGAC 1002  
 Qy 1121 AACAAATTTGTTTGGATCAACACTCAACCCAGCCTGTCTCAAGTACTGTGAGCAC 1180  
 Db 1003 AACAAGCTGTGTGTGAATCACTCGAGCTGAGCCCTGTGAGCAGATACAGGAGCAC 1062  
 Qy 1181 ACAGCAGCTGTAAAGTATTGATGCTCTCTCATCTTTCATGGAATCTTTCATCTGGA 1240  
 Db 1063 CTGGCGGCGCTGAAGGCAATCGCTGTCTCCCACTCAGCAGCGGCTGTGGCTCGGG 1122  
 Qy 1241 GAGGAACTGCAATGATGATGTTTGTGGAATACACCAACTCACACCTTAGC 1300

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 10:49:52 ; Search time 596 Seconds  
(without alignment)  
9085.692 Million cell updates/sec

Title: US-09-701-572-1  
Perfect score: 2006  
Sequence: 1 gatccgacaggaagaa.....agaaaaaaaaaaaaa 2006

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 19un03.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2006	100.0	2006	21	cdna encoding an a
2	339.2	16.9	5905	21	Human ORF3013
3	334.8	16.7	2645	23	Drosophila melanog
4	303.8	15.1	1794	24	Candida albicans e
5	269.6	13.4	3940	23	DNA encoding novel
6	263.8	13.2	1398	25	Aspergillus fumiga
7	263.8	13.2	1398	25	Aspergillus fumiga
8	263.8	13.2	1875	25	Aspergillus fumiga

9	263.8	13.2	1993	25	ABT20090
10	263.8	13.2	3398	25	ABT17680
11	263.8	13.2	3393	25	ABT17680
12	259	12.9	14283	23	ABL05492
13	248	12.4	1369	23	ABL18423
14	248	12.4	3369	23	ABL18422
15	248	12.4	5058	23	ABL05448
16	231.4	11.5	1706	23	ABL07411
17	213.6	10.6	1344	24	ABZ13183
18	213.2	10.6	3852	23	ABL07410
19	200.4	10.0	1767	16	AAQ96099
20	198.8	9.9	1686	24	ABT11002
21	198.8	9.9	1686	24	ABL65393
22	198.8	9.9	1686	24	ABL65837
23	198.8	9.9	1686	24	ABL67656
24	198.8	9.9	1686	24	ARK35539
25	198.8	9.9	1688	21	AAQ77801
26	198.8	9.9	1700	16	AAQ96100
27	191.4	9.5	2697	21	AAQ46256
28	174.8	8.7	650	21	AAF11864
29	157.2	7.8	14482	23	ABL10086
30	152.6	7.6	3404	23	ABL06722
31	147	7.3	274	25	ABX83461
32	132.2	6.6	574	23	ABX77748
33	131.6	6.6	645	24	ABN75375
34	113.4	5.7	1034	25	ABZ82891
35	109.6	5.5	480	24	ABT09540
36	109.6	5.5	672	24	ABL1366
37	101.2	5.0	416	21	AAH30586
38	97.6	4.9	484	24	ABL82189
39	92.8	4.6	1905	24	ABZ2237
40	92.6	4.6	457	24	ABL81347
41	90.8	4.5	469	24	ABZ35113
42	78.8	3.9	678	25	ABX7642
43	70	3.5	480	24	ABL81060
44	67.8	3.4	310	16	AAZ23988
45	66.2	3.3	486	24	ABN77135

ALIGNMENTS

RESULT 1  
AAZ37024  
ID AAZ37024 standard; cdna; 2006 BP.  
XX AC AAZ37024;  
XX DT 13-MAR-2000 (first entry)  
XX DE cdna encoding an alfalfa fzf protein designated ccc52Ms.  
XX KW Alfalfa; ccc52Ms; WD40 motif; fizzy-related; fzf subfamily; fzf protein;  
XX KW cellular differentiation; endoreplication; cell proliferation;  
XX KW plant cell; somatic embryogenesis; endoploidy; reserve tissue;  
XX KW in vitro plant regeneration; ss.  
XX OS Medicago sativa.  
XX FH Key  
XX CDS Location/Qualifiers  
FT Key  
FT CDS 182..1609  
FT FT /\*tag= a  
FT FT /product= "ccc52Ms"

FR2779433-A1.  
XX PD 10-DEC-1999.  
XX PF 08-JUN-1998; 98FR-0007174.  
XX PR 08-JUN-1998; 98FR-0007174.  
XX XX (CNRS ) CNRS CENT NAT RECH SCI.





QY 1681 TGAGCGGAGAACACCATGCTGGAAGAAACCTTGAATATATAAACAACCAACCAAGTAGCATC 1740  
AAC77458  
Db 1681 TGAGCGGAGAACACCATGCTGGAAGAAACCTTGAATATATAAACAACCAACCAAGTAGCATC 1740  
QY 1741 TTTACCAACTGGGAGAGCCTTGGAGGAGCTATATAAGTTTGTATATGGCTGGCGGTGAT 1800  
Db 1741 TTTACCAACTGGGAGAGCCTTGGAGGAGCTATATAAGTTTGTATATGGCTGGCGGTGAT 1800  
QY 1801 ATTCTGCAATCATGTGTAGTCTCATTTTATTTGAAAGATGATACAAATGGGTAATT 1860  
Db 1801 ATTCTGCAATCATGTGTAGTCTCATTTTATTTGAAAGATGATACAAATGGGTAATT 1860  
QY 1861 TATTCTCTTGGACTTATACATGATGAGTTGTAGCAAGTTTGTATATGGCTGGCGGTGAT 1920  
Db 1861 TATTCTCTTGGACTTATACATGATGAGTTGTAGCAAGTTTGTATATGGCTGGCGGTGAT 1920  
QY 1921 TTTTCTTCTCTTCTTTTGTATGATGCTCTCTCTGCTATTTATATATTTTAAATGCGT 1980  
Db 1921 TTTTCTTCTCTTCTTTTGTATGATGCTCTCTCTGCTATTTATATATTTTAAATGCGT 1980  
QY 1981 TTAACAGAGAAAAAAGAAAAA 2006  
Db 1981 TTAACAGAGAAAAAAGAAAAA 2006

RESULT 2  
AAC77458  
ID AAC77458 standard; cDNA; 5905 BP.  
XX AAC77458;  
AC AAC77458;  
DT 08-FEB-2001 (first entry)  
DE Human ORF3013 polynucleotide sequence SEQ ID NO: 6025.  
XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disease; diabetes mellitus; graft vs host disease;  
KW cardiovascular disease; systemic lupus erythematosus; infection;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200058473-A2.  
XX  
XX '05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000WO-US08621.  
XX  
XX 31-MAR-1999; 99US-0127607.  
XX 02-APR-1999; 99US-0127636.  
XX 05-APR-1999; 99US-0127728.  
XX 30-MAR-2000; 2000US-0540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Leach M;  
XX  
XX WPI; 2000-602362/57.  
XX P-PSDB; AAB43249.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 5206-5209; 5507pp; English.  
XX AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
CC antithyroid; and antianemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, diabetes mellitus,  
CC graft vs host disease, cardiovascular disease, osteoarthritis,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
SQ Sequence 5905 BP; 1095 A; 1770 C; 1963 G; 1076 T; 1 other;  
Query Match 16.9%; Score 339.2; DB 21; Length 5905;  
Best Local Similarity 62.7%; Pred. No. 1.8e-68;  
Matches 562; Conservative 0; Mismatches 328; Indels 6; Gaps 2;  
QY 647 CCTATATAGTTTGGATGCACCTGTTTGAAGATGATTTTATCTGAATCTGGTAGAT 706  
Db 784 CCCTTCAAGTGTGAGCGGCCCGAGCTGCAGGACGACTTCTACCTCAATCTGGTGAC 843  
QY 707 TGGTCTTACACAATGTGTGGCTGTGGTTGGGTAACGTGTCTATCTCTGGAATGCT 766  
Db 844 TGGTGTCTCCCTCAATGTCTCAGCTGGGGCTAGGCACTGCTGTACCTGTGGAGTGC 903  
QY 767 TGTAGCAGCAGGTACTAAATATATGTA---TTTGGGGCTTGTATGTTGTGTCT 823  
Db 904 TGTACACAGCAGGTGACCGGGCTCTGTGACCTCTCAGTGAAGGAGACTCAGTGACCTCC 963  
QY 824 GTTGGTTGGGCTCAACCGTGTACTCATCTTGTGTGGAACTAACTAAATGTAAGTTTCAG 883  
Db 964 GTGGGCTGGTCTGAGCGGGGAACTGTGTGGGGTGGGACACACAAAGGGCTTCGTGCAG 1023  
QY 884 ATTTGGGATCGACAGATCGAAGATGAAGATCAATGAGGGCCATCGTTCAGCTGTC 943  
Db 1024 ATCTGGGACGCGCCGAGGGAAGAGCTGTCCAATGTTGGAGGGCCACACGGACGGCTC 1083  
QY 944 GGGGCTTGGCTGGAGTTTCACTCTTTTGTCTTCTGTGGAGGGGATAGAATATTTAT 1003  
Db 1084 GGGGCTTGGCTGGAGTTTCACTGTGAGCAGCTGTGTGGGGAGCGCGACCATGATCCTG 1143  
QY 1004 CAACGAGATATACGCA---CACAAAGAGATTTTGTAGTAAACTGTGAGGACACAAATCA 1060  
Db 1144 CAGAGGAGCATCCGACACCCGCCCACTGCTGAGGCGGGCTGTCAGAGGGCCACCGGAG 1203  
QY 1061 GAGTTTGTGAGCTGAAGTGGTCAATATGATAACCGTGTGAGTTGGCATCTGGAGGAAATGAC 1120  
Db 1204 GAGTTGTGGGGCTCAGTGGTCCACAGACCCAGCTCCTCGCTCGGGGGCAACGAC 1263  
QY 1121 AACAAATTTTGTGTGGAAATCAACACTCAACCCAGCCTGTCTCTCAAGTACTGTGAGCAC 1180  
Db 1264 AACAAAGTCTGTGTGGAATCACTCGAGGCTGAGCCCCGTGTCAGAGTACACGAGGAC 1323  
QY 1181 ACAGCAGCTGTTAAAGCTATTGTCATGCTCTCATCTTGTGAGCTTCTTGTGATCTGGA 1240  
Db 1324 CTGGCGGCGGTGAAGGCAATCGCTGTCTCCCAATCAGACAGGGCTGTGGCTCGGG 1383  
QY 1241 GGAGGAATCGAGATAGATGATTTGGTTTGGAAATCAACCAAACTCACACTTAC 1300

Db 1384 GCGGCGACAGCTGACCGCTGTATCCGCTTCTGGRACACGCTGACAGGACCAACCACTGCAG 1443  
 QY 1301 TGTATGACACACTGGAAGTCAAGTTTGCAATCTTCTGTGCTCCAAAATGTCACAGAACTA 1360  
 Db 1444 TGTATGACACAGCGGCTCCCAAGTGTGCAATCTGCGCTGCTCCAAAGCAGCGCAAGAGCTG 1503  
 QY 1361 GTAAGCACACATGGGTACTCCAGAACACAGATTATTGTTTGGAGATACCCCACTATGTCA 1420  
 Db 1504 GTGAGCAGCAGCGCTACTCTCAGAACACAGATCTTGTCTGGAAGTACCCCTCCCTGACC 1563  
 QY 1421 AAGCTGGGACCTTTACCGGCCATCTATTAGGCTTCTATCTTGGCATCTCTCCAGAT 1480  
 Db 1564 CAGGTGGCCCAAGCTGACCGGGCACTCTACCGCGTGTGTACCTGGCAATGTCCCTGTAT 1623  
 QY 1481 GGACAGACTATTGTAACTGGAGCTGGAGATGAACGCTTAGGTTCTCGAAATGTTTT 1536  
 Db 1624 GGGGAGGCCATCGTCACTGGTGTGAGACAGAGACCCCTGAGGTTCTGGAAGCTTT 1679

## RESULT 3

ABL05493  
 ID ABL05493 standard; cDNA; 2645 BP.

XX ABL05493;  
 XX 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 10961.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEXE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR P-P8DB; ABB61390.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Claim 1; SEQ ID NO 10961; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16175-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (AB057737-AB072072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2645 BP; 766 A; 630 C; 619 G; 630 T; 0 other;

XX Query Match

XX Best Local Similarity 16.7%; Score 334.8; DB 23; Length 2645; E-value 1.5e-67;  
 XX Matches 572; Conservative 0; Mismatches 352; Indels 6; Gaps 2;

QY 614 CCGGTTAAGGCTCTCTAGGAAGGTTCTCTGATCCGCTTATAAGGTTTGAATGCACTGCT 673  
 Db 942 CCGGCGCAAGGCTACGCGCAAAATCTCTCGCAATTCCTTCAAGGTGTAGACGCGCCGAG 1001  
 QY 674 TTGCAAGATGATTTTATCTGGAATCTGGTAGATTGGTCTTTCACAAATGTTGGTGTGT 733  
 Db 1002 TTGCAAGACGACTTCTATCTGAACTGGTGTGCTGCTGCGAGAACGTTACTGGCTGTA 1061  
 QY 734 GGTTTGGGTAACCTGCTCTATCTCTGGAATGCTGTAGCAGCAAGTAATAATTATCT 793  
 Db 1062 GGCCTGGGAGCTGTGTCTATCTGTGGAGCGCTGTGACCACTCAGGTACCOCCTGTGT 1121  
 QY 794 GATTGGG---GGTTGATGATTGTGTTTGTGTTGTTGGGCTCAAGTGTGTAATCTAT 850  
 Db 1122 GATCTCAGTCCGATGCAATACGCTGACCTCGGTGTGCTGGAACGAGGCTGCAACACC 1181  
 QY 851 CTTGCTGTTGGAACCTAATGATGTAAGTTTCAATTTGGATGTCAGCAAGATGCAAGAG 910  
 Db 1182 GTGGCGGTGGGCAACATACCGGTACGTGACCGTCTGGGATGTGGCGCAATAAGCAG 1241  
 QY 911 ATAAGATCAATGGAGGCGCATCGTTACGTTGTCGGGCGCTTGGCCTGGAGTTTCATCT 970  
 Db 1242 ATCAACAACTGAATGGCCATTCGGCGCTGTGGCGCTTGGCATGGAACAGTGACATC 1301  
 QY 971 TTGCTTCTGCTGAGCGGATAGAATATTATCAACGAGATATACGCA---CACAAGAA 1027  
 Db 1302 CTGTGAGCGGTCGCGAGACCGTTGGATCATACGCGGATACGAGAACGCGCAACTG 1361  
 QY 1028 GATTTTGTTAGTAACTGTGAGACACAAATCAGAGGTTTGTGGACTGAGTGTGATAT 1087  
 Db 1362 CAATCGGAGCGCATGATTGGCGGACATCGCGAGAGGTGTGCGACTGAAATGGTCACT 1421  
 QY 1088 GATAACGCTGAGTTGGCATCTGGAGGAATGACAAATTTGTTCTTTGGATCAACAC 1147  
 Db 1422 GATTAATCAATCTTGGCCAGTGGCGCAACGATTAATCGGTTGATGTGGAATCAGCAT 1481  
 QY 1148 TCAACCCAGCCTGCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCAATG 1207  
 Db 1482 TCCGTGAATCCGTACAAATCATACAGGAGCATATGGCGCTGTAAAGGCGATCGCGTGG 1541  
 QY 1208 TCTCTCATCTTCATGACTTCTTGGCATCTGGAGGAGGAAGTGCAGATAGATGATTCGT 1267  
 Db 1542 TCGCGCATCAACACGAGCTCTTGGCAGCGCGGTGGAACGCGGATAGGTGATCCGT 1601  
 QY 1268 TTTTGAATACCAACCAAACTCACAACCTTAGCTGTATGAGCACTGGAAGTCAGGTTTC 1327  
 Db 1602 TTCTGGAATACGCTGACCGGCGCAGCCCATGCACTGGGTGGAACACGCGCTCGAGGTTTC 1661  
 QY 1328 AATCTGTGCTGGTCCAAAATGTCAGAACTAGTAGCAGACATGGGTACTCCCAAGAAC 1387  
 Db 1662 AATCTGCGCTGGTCCAAGCACTCTCGGAGCTGCTCTCCAGCACGCGCTACTTCGAGAAC 1721  
 QY 1388 CAGATTATTGTTGGAGATACCCCACTATGTCAAGCTGGGCACTCTTACCAGCCATATCT 1447  
 Db 1722 CAGATCTCTGTGGAATATCCCTCCCTGACCAAGTGGCCAAAGTACGCGGCCATTCTG 1781  
 QY 1448 TATAGGTTCTCTATCTTGGCATCTCTCCAGATGGAAGCACTATTCTTAATCGAGCTGGA 1507  
 Db 1782 TATCGTGTGCTCTATCTGCGGCTGAGTCCGATGGTGAAGCTATTGTTACGGCGCGC 1841  
 QY 1508 GATCAAAACGCTAGGTTCTCGAATGTTTC 1537  
 Db 1842 GACGAGACGCTGCGATTGTTGGAACGTATTC 1871

## RESULT 4

ABZ32308  
 ID ABZ32308 standard; DNA; 1794 BP.

XX AC ABZ32308;

XX DT 30-JAN-2003 (first entry)



XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX PR 31-MAR-2000; 2000US-0540217.  
XX PR 23-AUG-2000; 2000US-0649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Dmanac RT, Liu C, Tang YT;  
XX DR WPI; 2001-639362/73.  
XX DR P-PSDB; ABG13562.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity -  
XX PS Claim 1; SEQ ID No 13553; 103pp; English.  
XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 3940 BP; 799 A; 1246 C; 1221 G; 674 T; 0 other;  
Query Match 13.4%; Score 269.6; DB 23; Length 3940;  
Best Local Similarity 62.6%; Pred. No. 2.1e-52;  
Matches 438; Conservative 0; Mismatches 259; Indels 3; Gaps 1;  
QY 840 GTGGTACTCATCTTCTGCTGGAACATAAATGTAAGTTCAGATTGGATGACGAA 899  
DB 2357 GGGGGAACCTGGTGGCGGTGGGACACACAAAGGGCTTCGTGACAGATCTGGGACGCG 2416  
QY 900 GATGCAAGATATAGATCAATGAGGGCCATCGTTACGTGCGGGCTTGGCCCTGGA 959  
DB 2417 CAGGGAAGAAGTGTCATGTTGGAGGGCCACACGGCAGCGTTCGGGCGCTGGCTTGA 2476  
QY 960 GTTCATCTCTTTTCTTCTGTCGACGGGATAGAAATATTATCAACAGATATACGA 1019  
DB 2477 ATGCTGAGCAGCTGCTGCTCGGGAGCGCGGACCGCATGATCTCTGACAGGAGCATCGCA 2536  
QY 1020 ----CACAGAAGATTTTGTAGTAACTGTACGACACAAATCAGAGTTTGTGACTGA 1076  
DB 2537 CCCCAGCACTGAGTCGAGCGGGCGGTGACGGGGCCACCGGAGAGGTGTGGGGCTCA 2596  
QY 1077 AGTGGTCAATATCAATACCGTGAAGTTGGCATCTGGAGGAATGACAAATTTGTTGTT 1136  
DB 2597 AGTGGTCCACAGACCAGCTCTCGCTCGGGGGGCAACGACAAAGCTGTGGTGTCT 2656  
QY 1137 GGAATCAACACTCAACCCAGCTGCTCTCAAGTACTGTGAGCACACACAGCTGTAAAG 1196  
DB 2657 GGAATCACTCAGAGCTGAGCCCGGTGACAGCAGTACACGGAGCACCTGGCGGCGGTAAGG 2716

QY 1197 CTATTGATGGTCTCCCTCATCTTCATGGACTTCTTGTGATCTGGAGGAGGAACTCGAGATA 1256  
DB 2717 CCATCGCTGCTCCCCACATCAGCAGCGGTGCTGGCTCGGGGGGGCGGACAGCTGACC 2776  
QY 1257 GATGATTTGGTTTGGAAATACAAACACAAACTCACAACCTTAGCTGTATGGACACTGGAA 1316  
DB 2777 GCTGTATCCGCTTCTGGACACACGCTGACAGGACAACTGCACTGTATGACACAGCGCT 2836  
QY 1317 GTCAGGTTTGCATCTTGTCTGGTCCAAAATGTCAACGAAGTGTAGTAAGCACACATGGT 1376  
DB 2837 CCCAAGTGTGCAATCTGGCTGCTGCAAGCAGCGCAAGAGCTGGTGAGCAGCGACGCT 2896  
QY 1377 ACTCCAGAACAGATATTATTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTA 1436  
DB 2897 ACTACAGAACAGATCTTGTCTGGAAGTACCCCTCCCTGACCCAGCTGGCCAACTGA 2956  
QY 1437 CCGGCAATCTTATGAGGTTCTTATCTTGCCATCTCTCCAGATGGACAGACTATTCTAA 1496  
DB 2957 CCGGCACTCTTACCGGTGCTGTACCTGCGCAATGTCCCTGATGGGAGGCCATCTCA 3016  
QY 1497 CTGGAGCTGGAGATGAACAGCTTAGGTTCTGGAATGTTT 1536  
DB 3017 CTGGTCTGGAGACGAGACCCCTGAGTTCTGGAAGCTCTT 3056  
RESULT 6  
ABT18274  
ID ABT18274 standard; DNA; 1398 BP.  
XX AC ABT18274;  
XX DT 16-APR-2003 (first entry)  
XX DE Aspergillus fumigatus essential gene #632.  
XX KW Fungicide; cytosolic; essential gene; Aspergillus fumigatus; infection;  
XX KW cancer; contamination; biofilm; antibody; immune response; ds.  
XX OS Aspergillus fumigatus.  
XX PN WO200286090-A2.  
XX PD 31-OCT-2002.  
XX PF 23-APR-2002; 2002WO-US13142.  
XX PR 23-APR-2001; 2001US-285697P.  
XX PR 27-APR-2001; 2001US-287066P.  
XX PR 05-JUN-2001; 2001US-295890P.  
XX PR 09-JUL-2001; 2001US-303899P.  
XX PR 31-AUG-2001; 2001US-316362P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;  
XX WPI; 2003-093124/08.  
XX DR New purified or isolated nucleic acids of essential genes of  
XX PT Aspergillus fumigatus, useful for treating or preventing infections by  
XX PT A. fumigatus, or for treating a non-infectious disease in a subject  
XX PT e.g. cancer -  
XX XS Disclosure; Page -; 175pp; English.  
XX CC The invention relates to novel purified or isolated nucleic acids of  
XX CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
XX CC the invention are used to treat or prevent infections by a pathogenic  
XX CC organism such as A. fumigatus, to treat a non-infectious disease in a  
XX CC subject (e.g. cancer), to prevent or contain contamination of an object  
XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for

expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This polynucleotide sequence represents one of the essential genes of *Aspergillus fumigatus* of the invention.

XX  
SQ Sequence 1398 BP; 368 A; 360 C; 343 G; 327 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1398;  
Best Local Similarity 56.3%; Pred. No. 3.4e-51;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCCTAGGAAGGTTCTCGATCCCTTATAAGGTTTGGATGCAC 668  
DB 341 AGACTCCTCGCAAGCAGCCTCGTACGTTAATAAAGTACCTTATAAGGTTCTCGACGCAC 400

QY 669 CTGCTTTCAGATGATTTTATCTGATCTGATGATGCTTCACACATGTTGG 728  
DB 401 CGGATTTGCAAGATGATTTCTACTTGAATCTGCTGATCGGCGAGTAGTAAATGCTAG 460

QY 729 CTGTTGGTTTGGGTAACTGTGCTATCTCTGGAATGCTTGTAGCAGCAAGTAATAAT 788  
DB 461 GGTGTTGGCTAGCAATTCAGTGTACATGTGGAATTCGAATACCGACGGGTGACGAAC 520

QY 789 TATGTGATTTGGGGTTGATGTTG-----TGTTTGTCTGTTGGTCTCAAGTGTGCTC 813  
DB 521 TTTCGCACTAAGAGATGATGATGTCACAAAGTTAGTGTGATTCAGAGGTTAAGAT 580

QY 814 -----TGTTTGTCTGTTGGTCTCAAGTGTGCTC 848  
DB 581 TTCTCTATATGTTGGCAATGATTAATATGTTTCTAGCTGACATTCGACAGGSCACAC 640

QY 849 ATCTGCTGTTGGAACCTAACAATGTTAAAGTTTCAATTTGGATTCGACAGAGATCAAGA 908  
DB 641 ATCTTTCAATAGGAACCTGCAAAAGGTTTGTACAGATATGGGATCGAGCATTTGCGCC 700

QY 909 AGATAAGATCAATGGAGGCGCCATCGTTACGTGTGGGGCCCTTGGCTGGAGTTTCATCTC 968  
DB 701 GTCTTCGGAACATGATTTGGGATACCAATCGTGTGGGGCTCTTCTTGGACGATCATA 760

QY 969 TTTTGTCTTCTGGTGGACGGGATAGAATATTTATCAACGAGATATACGACACAAGAG 1028  
DB 761 TCTTACGTCAGGTTCTCGGATCGACTGATCTTCCACCGTGTGTTGTTCCCGCAGATC 820

QY 1029 ATTTGTTAGTAACTGTGACGACCAATCAGAGTTTGTGCACTGAAGTGTGTCATATG 1088  
DB 821 AGTATTTACGTCAGTCTCGGCGCCATACGACGAAGTTTCGGATCTCGGTGGAACACCG 880

QY 1089 ATAACCGTGAAGTTGGCATCTGAGGAAATGACAACAATTTGTTTGGATCAACACT 1148  
DB 881 AAGATGCGCACTAGCTTCAGGCGGAATGACAACAATCTCATGTGTGGGACAAGTGA 940

QY 1149 CAAACCGGCTGCTCAAGTACTGTGACACACAGCAGCTTTAAAGCTTATGATGCT 1208  
DB 941 ATGAGACACCGCTTTATCGCTTCTCGGACCATACCGACCGGTGAAGGCCATCGCTGCT 1000

QY 1209 CTCTCATCTTCAATGACTTCTTGCACTCTGGAGGAGAACTGCAGATAGATGATTCGTT 1268  
DB 1001 CACCCCATCAACACCATCTCTCTCGCTTCAGGTGGAGGACGCGGACCGACCATCAAT 1060

QY 1269 TTTGGAATACACCAAACTCACACCTTAGCTGTATGGACACTGGAAGTCAGGTTTGCA 1328  
DB 1061 TCTGGACACTGCTACTGTTTCACTGATCAAGAAAGTTGATACGGTAGCAGGTGTGTA 1120

QY 1329 ATCTTCTCTGGTCCAAAATGTCAAGAACTAGTAAGCACACATGGTACTCCAGAAC 1388  
DB 1121 ACCGTGCATGTGCGAAGAACTCGGACGAATCATCAGTACACACCGCTACATCAGAACC 1180

QY 1389 AGATTATTTTGGAGATACCCCACTATCTCAAAGCTGGCGACTCTTACCGGCCATACTT 1448  
DB 1181 AGATTGTAATCTGGAATACCCGCGCATGGACAGATTGTATCGCTGACCGGCCACACTT 1240

QY 1449 ATAGGTTCTCTATCTTTCGCAATCTCTCCAGATGGACAGACTATTTAATCTGAGTGGAG 1508  
DB 1241 TCCGTGTTCTTACTCTTGGATGAGCCCGATGGACAGACCGTCTGTTACCGGAGTGGTG 1300

QY 1509 ATGAAACGCTTAGTCTCGAATGTTTTTTC 1537  
DB 1301 ACGAAACGTTGAGATTCTGGAAGATTTTC 1329

## RESULT 7

ABT18868

ID ABT18868 standard; DNA; 1398 BP.

XX ABT18868;

DT 16-APR-2003 (first entry)

DE *Aspergillus fumigatus* essential gene #1226.XX Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection; cancer; contamination; biofilm; antibody; immune response; ds.XX *Aspergillus fumigatus*.

XX WO200286090-A2.

XX 31-OCT-2002.

XX 23-APR-2002; 2002WO-US13142.

XX 23-APR-2001; 2001US-285697P.

XX 27-APR-2001; 2001US-287066P.

XX 05-JUN-2001; 2001US-295890P.

XX 09-JUL-2001; 2001US-303899P.

XX 31-AUG-2001; 2001US-316362P.

XX (ELIT-) ELITRA PHARM INC.

XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of

XX *Aspergillus fumigatus*, useful for treating or preventing infections by

XX A. fumigatus, or for treating a non-infectious disease in a subject

XX e.g. cancer

XX Disclosure; Page -: 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biofilm comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or



CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This polynucleotide sequence represents one of the essential  
CC genes of *Aspergillus fumigatus* of the invention.  
XX  
SQ Sequence 1398 BP; 368 A; 360 C; 343 G; 327 T; 0 other;  
  
Query Match 13.2%; Score 263.8; DB 25; Length 1398;  
Best Local Similarity 56.3%; Pred. No. 3.4e-51;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;  
  
QY 609 ATAGTCGGTAAAGGCTCCTAGGAAGTTCCTCGATCGCCCTTATAAGGTTTGGATGCAC 668  
DB |||||  
QY 341 AGACTCCTCGCAACAGCCCTCGTACGTTAATAAAGTACCCTATAAGGTTCTCGACGCAC 400  
DB |||||  
QY 669 CTCCTTTGCAAGATGATTTTATCTCAATCTGATAGTTGGTCTTCACACAATGTTGG 728  
DB |||||  
QY 401 CCGATTTGCAAGATGATTTTCTACTTGAATCTGGTGCATGGGCGAGTAGTATGTCAG 460  
DB |||||  
QY 729 CTGTTGGTTGGTAACTGTTCTATCTCTGGAATCTGTTAGCAGCAAGGTAACTAAT 788  
DB |||||  
QY 461 GTGTTGGCTTAGGCAATTCAGTGTACATGTTGAATTCGAATACCGGACCGGTTGACGAAC 520  
DB |||||  
QY 789 TATGTGATTTGGGGTTGATGTTG----- 813  
DB |||||  
QY 521 TTTGCGAATAAGATGATGATCTGCAAGTGTAGTGGATTACAGAGGTAATGAGAT 580  
DB |||||  
QY 814 -----TGTTTGTCTCTGTTGGTGGCTCAACGTGGTACTC 848  
DB |||||  
QY 581 TTCTCTATATGTTGGCAATGATTTATGTTTCTGATCTGACATTTGGACAGGGCACAC 640  
DB |||||  
QY 849 ATCTTGCTGTGGAATCAACAAAGTAAAGTTTGGATGTCAGCAAGATGCAAGA 908  
DB |||||  
QY 641 ATCTTTCAATAGGAAGTGGCAAGGCTTGTACAGATATGGATGCAAGCATTTGCGCC 700  
DB |||||  
QY 909 AGATAAGATCAATGAGAGGCGCATCGGTTACGTCTCGGGGCTTGGCTGGAGTTCACTC 968  
DB |||||  
QY 701 GTCTTCGCAATGATTTGGCATACCAATCGTGTGGGGCTCTTGTGGAAAGATCATA 760  
DB |||||  
QY 969 TTTTGTCTTCTGGTGGACGGGATAAGATATTTATCAACGAGATATACGACACAGAAG 1028  
DB |||||  
QY 761 TCCTTACGTAGGTTCTCGGGATCGACTGATCTCCACCGGTGATGTTGTTCCCGAGATC 820  
DB |||||  
QY 1029 ATTTTGTAGTAAACTGTGAGGACACAAATCAGAGGTTTGTGACTGAAAGTGTCTATG 1088  
DB |||||  
QY 821 AGTATTACGTGCACTCTCGGCCATTAAGCAGGAAGTTTCGGGACTCGGTGAACACCG 880  
DB |||||  
QY 1089 ATAACGTTGATTTGGCACTGAGGAATGACAAATTTGTTTGGTGAATCAACACT 1148  
DB |||||  
QY 881 AAGATGGCCAACTAGTCTTCAAGCGGAAATGACAAACTCATGGTGTGGGACAACTGA 940  
DB |||||  
QY 1149 CAACCCAGGCTGCTCCTCAAGTACTGTGAGCAGACAGCAGCTGTTAAAGCTATTGTCATGGT 1208  
DB |||||  
QY 941 ATGAGACACCGCTTATTCGCTTCTCGGACCATACCGGCGGTGAAGCCATCGCTGCT 1000  
DB |||||  
QY 1209 CTCCTCATTTTATGGAATCTTTCGATCTGAGGAGGAATCGCAGATAGATGATTCGTT 1268  
DB |||||  
QY 1001 CACCCCATCAACACCATCTCTCGCTCCTCAGTGGAGGACCGGCGGACCGACCATCAAT 1060  
DB |||||  
QY 1269 TTTGGATACACCAACCAACTCACCTTACCTGTATGGAACACTGGAAGTCAAGTGTGCA 1328  
DB |||||  
QY 1061 TCTGAAACATGCTACTGGTTCATCTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1120  
DB |||||

QY 1329 ATCTTGTCTGGTCCRAAATGTCAAGAACTAGTAGCAGACATGGGTACTCCAGAAC 1388  
DB |||||  
QY 1121 ACCTGGCATGGTCCGAAGAACTCGGAGGAATCATCTAGTACACCGCTACGTCAGAAC 1180  
DB |||||  
QY 1389 AGATTATTGTTTGGAGATACCCCACTATCTCAAAAGCTGGGACTCTTACCGGCCATACTT 1448  
DB |||||  
QY 1181 AGATTGTAATCTGGAATACCCCGCATCGAGCAGATTCTATCGCTGACCGGCCACACTT 1240  
DB |||||  
QY 1449 ATAGGCTTCTCTATCTGTCATCTCTCCAGATGGACAGACTATTGTAATCTGGAGCTGGAG 1508  
DB |||||  
QY 1241 TCCGTGTTCTCTACCTTGGCATGAGCCCGATGAGCCGATGAGCAGACCCGTCGTTACCGGAGCTGGTG 1300  
DB |||||  
QY 1509 ATGAAACGCTTAGGTTCTGGAATGTTTC 1537  
DB |||||  
QY 1301 ACGAACGTTGAGATTCTGGAAGATTTTC 1329  
DB |||||  
  
RESULT 8  
ABT20688  
ID ABT20688 standard; DNA; 1875 BP.  
XX  
AC ABT20688;  
DT 16-APR-2003 (first entry)  
XX  
DE *Aspergillus fumigatus* essential gene #3046.  
XX  
KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
KW cancer; contamination; biofilm; antibody; immune response; ds.  
XX  
OS *Aspergillus fumigatus*.  
XX  
FN WO200286090-A2.  
XX  
PD 31-OCT-2002.  
XX  
PF 23-APR-2002; 2002WO-US13142.  
XX  
PR 23-APR-2001; 2001US-285697P.  
PR 27-APR-2001; 2001US-287066P.  
PR 05-JUN-2001; 2001US-295890P.  
PR 09-JUL-2001; 2001US-303899P.  
PR 31-AUG-2001; 2001US-316362P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;  
XX WPI; 2003-093124/08.  
XX  
DR New purified or isolated nucleic acids of essential genes of  
PT *Aspergillus fumigatus*, useful for treating or preventing infections by  
PT *A. fumigatus*, or for treating a non-infectious disease in a subject  
PT e.g. cancer -  
XX  
XX Disclosure; Page -; 175pp; English.  
XX  
CC The invention relates to novel purified or isolated nucleic acids of  
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
CC the invention are used to treat or prevent infections by a pathogenic  
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a  
CC subject (e.g. cancer), to prevent or contain contamination of an object  
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a  
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
CC expressing recombinant protein for characterization, screening or  
CC therapeutic use, as markers for host tissues in which the pathogenic  
CC organisms invade or reside, for comparing with the DNA sequence of *A.*  
CC *fumigatus* to identify duplicated genes or paralogues having the same or  
CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination



CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This polynucleotide sequence represents one of the essential  
CC genes of *Aspergillus fumigatus* of the invention.  
XX

SQ Sequence 1875 BP; 514 A; 490 C; 464 G; 407 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1875;  
Best Local Similarity 56.3%; Pred. No. 3.7e-51;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCTCAGGAAGTTCTTCGATCGCTTATAAGGTTTGGATGCAC 668  
DB 818 AGACTCTCGAAGCAGCTCGCTAGCTTATAAAGTACCCTATAAGGTTCTCGACGCAC 877  
QY 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGCTAGATGCTTTCACACATGTTGG 728  
DB 878 CCGATTTCGAAGATGATTTTCTACTTGAATCTGCTAGCTGCGGCACTAGTAAAGTCTAG 937  
QY 729 CTGTTGGTTGGTAACTGCTCTATCTCTGGAATGCTTTGACGACGAAGGTAACATAAT 788  
DB 938 GTGTTGGCTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGGACGGGTGACGAAC 997  
QY 789 TATGATTTGGGGTGTGATGTTG----- 813  
DB 998 TTTGCGAATTAAGATGATGATCTGTCAAAAGTTAGTGGATTCAGAGGGTAATGAGAT 1057  
QY 814 -----TGTGTTGTTCTGTTGTTGGGCTCAACGTGTGATCTC 848  
DB 1058 TTCTCTATATGTTGGCAATGATTAATATGTTTCATGACTGACATTTGGACAGGCACAC 1117  
QY 849 ATCTGCTGTGGAATCAATGTAAGTTGAGTTGGATGCGACGACAGATGCAAGA 908  
DB 1118 ATCTTTCAATAGGAACCTGGCAAGGCTTTGTACAGATATGGGATGCGAGCAATTCGCGC 1177  
QY 909 AGATAAGATCAATGGAGGCGCATCGGTACGTGTGCGGCGCTTGGCTGGAGTTTATCTC 968  
DB 1178 GTCTTCGACATATGATTTGGCATACCAATCGTGTGGGGCTCTTGTGGACGATCAT 1237  
QY 969 TTTTGTCTTCTGTGACCGGATTAAGAAATATTATTCACAGAGATATACGCAACAAGAG 1028  
DB 1238 TCCTTACGTGAGTTCTCGGATCGACTGATCTTCCACCGTGATGTTCTCCCGAGATC 1297  
QY 1029 ATTTTGTAGTAACTGTGAGGACACAATTCAGAGCTTTGTGACTGAAAGTGTCTATATG 1088  
DB 1298 AGTATTTTACGTGACTGTCCGGCCATAGCAGGAAGTTTGGCACTCCGGTGGACACCG 1357  
QY 1089 ATAACCGTGTAGTTGGCATCTGGAGGAATGACAAACAAATTTGTTTGGATCAACACT 1148  
DB 1358 AAGATGCCCACTAGCTTTCAGGCGGAATGACAAACAACTCATGTTGGTGGGCAAGCTGA 1417  
QY 1149 CAACCCAGCTGTCTCAAGTATCTGTGACACACAGCAGTGTAAAGTATTCATGCT 1208  
DB 1418 ATGAGACACCGCTTTATTCGTTCTTCGACCATACCGACCGCTGAAAGGCACTCGCGTGT 1477  
QY 1209 CTCTCATCTTCATGGAATCTTGTGCATCTCGAGGAGAACTGACAGATAGATGTTTCGTT 1268  
DB 1478 CACCCCATCAACACCATCTCTCGCTCAGTGTGAGGCGCGGCGGACCGACCATCAAT 1537  
QY 1269 TTTGGAATACACCAACATCACTCATCTAGCTGTATGACATCTGGAATCTGAGTTTGA 1328  
DB 1538 TCTGGAACACTGTACTGTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1597  
QY 1329 ATCTTGTCTGTCCAAAATGTCACCAACTAGTAAGCACACATGTTGATCTCCCAAGAC 1388  
DB 1598 ACCTGGCATGTTCGAAGAACTCGGACGAATCATCAGTACACACGCTACATCAGAAC 1657

QY 1389 AGATTATTGTTGGAGATACCCACACTATGTCAAAGCTGGCGACTCTTACGGCCACTT 1448  
DB 1658 AGATTGTAATCTGGAATATACCCCGCATGGACGAGATTGTATCGTACCGGCCACACT 1717  
QY 1449 ATAGGTTTCTTCTATCTTGGCATCTCTCCAGATGACAGACTATTGTAACTGGAGCTGGAG 1508  
DB 1718 TCCGTGTTCTTCTACTACCTTGGCATGAGCCCGCATGACAGACCCGTCGTTACCGGAGCTGGTG 1777  
QY 1509 ATGAAGAGCTTAGTCTCTGGAATGTTTC 1537  
DB 1778 ACGAACGTTGAGATTCTGGAAGATTTC 1806

RESULT 9

ABT20090  
ID ABT20090 standard; DNA; 1993 BP.

XX  
AC ABT20090;

XX  
DT 16-APR-2003 (first entry)

XX  
DE *Aspergillus fumigatus* essential gene #2448.

XX  
KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
cancer; contamination; biofilm; antibody; immune response; ds.

XX  
OS *Aspergillus fumigatus*.

XX  
PN WO200286090-A2.

XX  
PD 31-OCT-2002.

XX  
PF 23-APR-2002; 2002WO-US13142.

XX  
PR 23-APR-2001; 2001US-285697P.

XX  
PR 27-APR-2001; 2001US-287066P.

XX  
PR 05-JUN-2001; 2001US-295890P.

XX  
PR 09-JUL-2001; 2001US-303899P.

XX  
PR 31-AUG-2001; 2001US-316362P.

XX  
PA (ELIT-) ELITRA PHARM INC.

XX  
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX  
DR WPI; 2003-093124/08.

XX  
PT New purified or isolated nucleic acids of essential genes of  
*Aspergillus fumigatus*, useful for treating or preventing infections by

XX  
PT *A. fumigatus*, or for treating a non-infectious disease in a subject  
e.g. cancer

XX  
PS Disclosure; Page -; 175pp; English.

XX  
CC The invention relates to novel purified or isolated nucleic acids of  
essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
the invention are used to treat or prevent infections by a pathogenic  
organism such as *A. fumigatus*, to treat a non-infectious disease in a  
subject (e.g. cancer), to prevent or contain contamination of an object  
by *A. fumigatus*, or to prevent or inhibit formation on a surface of a  
biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
expressing recombinant protein for characterization, screening or  
therapeutic use, as markers for host tissues in which the pathogenic  
organisms invade or reside, for comparing with the DNA sequence of *A.*  
*fumigatus* to identify duplicated genes or paralogues having the same or  
similar biochemical activity and/or function, for comparing with DNA  
sequences of other related or distant pathogenic organisms to identify  
potential orthologous essential or virulence genes, for selecting and  
making oligomers for attachment to a nucleic acid array for examination  
of expression patterns, for raising anti-protein antibodies, as an  
antigen to raise anti-DNA antibodies or to elicit another immune  
response, and for identifying polynucleotides encoding the other protein  
with which binding occurs or to identify inhibitors of the binding

CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This polynucleotide sequence represents one of the essential  
CC genes of Aspergillus fumigatus of the invention.

XX  
SQ Sequence 1993 BP; 543 A; 516 C; 496 G; 438 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1993;  
Best Local Similarity 56.3%; Pred. No. 3.8e-51;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

Qy 609 ATAGTCGGTTAGGCTCCAGGAGTTCTCGATCGCTTATAGGTTTGGATGCAC 668  
Db 936 AGACTCTCGCAGACGCTCGTACGTTAATAAGTACCCTTAAAGTTCTCGACGAC 995  
Qy 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGTTAGATGGTCTTCACAAATGTTGG 728  
Db 996 CCGATTTGCAAGATGATTTTCTGATCTGTTGAGCTGGGCGAGTAGTAAATGCTAG 1055  
Qy 729 CTGTTGTTGGTAACTGTTCTATCTCTGGAATGCTTGTACGACGAAGTAACAT 788  
Db 1056 GTGTTGCCCTAGGCAATTCAGTGTACATGTGGAATTCGAATCCGACGGGTGACGAAC 1115  
Qy 789 TATGTGATTTGGGGGTTGATGATTG----- 813  
Db 1116 TTTGCGAAGTAAAGATGATGACTGTCAAGTGTAGCTGATTCAGAGGTAATGAGAT 1175  
Qy 814 -----TGTGTTGTTCTGTTGGTCTCAAGTGGTACTC 848  
Db 1176 TTTCTTATATGTTGGCAATGATTTATGTTTCTGATGACTGACATTTGGACGGGACAC 1235  
Qy 849 ATCTGCTGTTGGAACTAACTAAGTAAAGTTCAGATTGGGATGACGAAGATGCAAGA 908  
Db 1236 ATCTTTCAATAGAACTGCGCAAGTCTTGTACAGATATGGATGCGAGCATTTGCGCC 1295  
Qy 909 AGATAAGATCAATGGAGGCGCATCGGTTACGTTGCGGGGCTTGGCTGGAGTTCAATCT 968  
Db 1296 GTCTTCGGACAAATGATTTGGGCAATACCAATCGTGTGGGGGCTTGTCTTGGACGATCAT 1355  
Qy 969 TTTTGTCTTCTGTTGGAGCGGATGAAGATATTTATCAAGAGATATACGACACAAGAAG 1028  
Db 1356 TCCTTAGCTAGGTTCTCGGATCGACTGATCTCCACCTGATGTTGCTGCTCCGAGATC 1415  
Qy 1029 ATTTTGTAGTAACTCTGACGACACAATAATCAGAGGTTTGTGACTGAAGTGTCTATG 1088  
Db 1416 AGTATTTACGTCGACTGTCCGGCCATTAAGCAGGAAGTTTGGGACTCCGGTGAACACCG 1475  
Qy 1089 ATACCGTGTAGTTGGCATCTGGAGGAATGACAAATGTTGTTGTTGAATCAACACT 1148  
Db 1476 AAGATGGCCAACTAGCTTCAGGCGGAATGACAACTCATGTTGTTGGACAAGCTGA 1535  
Qy 1149 CAACCCAGCTGCTCCTCAAGTACTGTGAGCAACAGCAGCTGTTTAAAGCTATTGCAATGGT 1208  
Db 1536 ATGAGACACCGCTTTATCGCTTCTCGGACCATATCCGACGCGTGAAGCCATCGCTGGT 1595  
Qy 1209 CTCTCATCTTCATGGACTTCTTCATCTGAGAGGAACTGACAGATAGTATGTTGTT 1268  
Db 1596 CACCCCATCAACACATCTCTCGCTCAGTGTGAGGACGCGGACCGGACCATCAAT 1655  
Qy 1269 TTTGGAATACACACAACTCACAACCTTAGCTGTATGGACATCTGGAAGTCAAGTTTGCA 1328  
Db 1656 TCTGGACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1715  
Qy 1329 ATCTTGTCTGTGCGAATAATGTCACGAACTAGTAGACACACATGGGTACTCCAGAAC 1388  
Db 1716 ACCTGGCATGTGCGAAGAACTCGGACGAATCATCAGTACACACGGGTACAGTCAGAAC 1775  
Qy 1389 AGATTATTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTTACCGGCCCACT 1448  
Db 1776 AGATTGTAACTGGAATACCGCGCATCGGACGAGAGATGTTATCGCTGACCGGCCACACT 1835

Qy 1449 ATAGGTTCTTATCTTCCATCTCTCCAGATGACAGACTATTTGTAACCTGGAGCTGAG 1508  
Db 1836 TCCGTGTTCTTACTTCCGATGAGCCCGATGACAGACCGTGTACCGAGCTGGT 1895  
Qy 1509 ATGAAAGCTTGTAGTTCTTGAATGTTTTC 1537  
Db 1896 ACGAACGTTGAGATTCGGAAGATTTTC 1924

RESULT 10  
ABT17680  
ID ABT17680 standard; DNA; 3398 BP.  
XX  
AC ABT17680;  
XX  
DT 16-APR-2003 (first entry)  
XX  
DE Aspergillus fumigatus essential gene #39.  
XX  
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
KW cancer; contamination; biofilm; antibody; immune response; ds.  
XX  
OS Aspergillus fumigatus.  
XX  
PN WO200286090-A2.  
XX  
PD 31-OCT-2002.  
XX  
PP 23-APR-2002; 2002WO-US13142.  
XX  
PR 23-APR-2001; 2001US-285697P.  
PR 27-APR-2001; 2001US-287066P.  
PR 05-JUN-2001; 2001US-295890P.  
PR 09-JUL-2001; 2001US-303899P.  
PR 31-AUG-2001; 2001US-316362P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;  
XX  
DR WPI; 2003-093124/08.  
XX  
PT New purified or isolated nucleic acids of essential genes of  
PT Aspergillus fumigatus, useful for treating or preventing infections by  
PT A. fumigatus, or for treating a non-infectious disease in a subject  
PT e.g. cancer  
XX  
PS Disclosure; Page -; 175pp; English.  
XX  
CC The invention relates to novel purified or isolated nucleic acids of  
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
CC the invention are used to treat or prevent infections by a pathogenic  
CC organism such as A. fumigatus, to treat a non-infectious disease in a  
CC subject (e.g. cancer), to prevent or inhibit formation of an object  
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
CC expressing recombinant protein for characterisation, screening or  
CC therapeutic use, as markers for host tissues in which the pathogenic  
CC organisms invade or reside, for comparing with the DNA sequence of A.  
CC fumigatus to identify duplicated genes or paralogues having the same or  
CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to

CC isolate correlative receptors or ligands in the case or virulence  
CC factors. This polynucleotide sequence represents one of the essential  
CC genes of Aspergillus fumigatus of the invention.  
XX

SQ Sequence 3398 BP; 884 A; 845 C; 789 G; 880 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 3398;  
Best Local Similarity 56.3%; Pred. No. 4.5e-51;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

```
QY 609 ATAGTCGGTTAAGGCTCTAGAGGTTCCGATCGCTTATAAGGTTTGGATGCAC 668
DB 1341 AGACTCTCGCAGCAGCCCTCAGTTAATAAAGTACCCCTATAAGGTTCTCGACGAC 1400
QY 669 CTGCTTTGCAAGATGTTTTTATCTGAATCTGGTAGATTGGTCTTCACACAAATGTTGG 728
DB 1401 CCGATTTGCAAGATGTTTTTATCTGAATCTGGTAGATTGGTCTTCACACAAATGTTGG 1460
QY 729 CTGCTTTGTTGGTAACTGTTCTATCTCTGAAATGCTTTGAGCAGCAAGTTAACTAAT 788
DB 1461 GTTGTGGCTTAGGCAATTCAGTGTACATGTGGAATTCGAAATACCGGACGGGTGACGAAAC 1520
QY 789 TATGTGATTTGGGGTGTGATGTTG----- 813
DB 1521 TTTGCGAATAGAGATGATCTGTCAAGTGTAGCTGGATTGAGGGTAAATGAGAT 1580
QY 814 -----TGTGTGTCTGTTGGTGGCTCAACGTGTACTC 848
DB 1581 TTCTCTATATGTTGGCAATGATTAATATGTTTTCATGACTGACATTTGGACAGGGCACAC 1640
QY 849 ATCTTCTCTGTTGAACTAACAATGTAAGTTTTCAGATTTGGATGTCAGCAGCATCAAGA 908
DB 1641 ATCTTTCAATAGGAATGGAAGGCTTTGTACAGATATGGATGCGAGCATGTTGCGC 1700
QY 909 AGATAAGATCAATGAGGCGCATCGGTTACGTTGCGGCGCTTGGCTGGAGTTCACTC 968
DB 1701 GTCTTCGGAATGATTTGGGCAATACCAATCGTTGGGGGCTCTGTCTTGGAACTGATA 1760
QY 969 TTTTGTCTTCTGTTGAGCGGATGAAGATATTTATCAAGAGATATACGACACAGAGAG 1028
DB 1761 TCCTTACGTGAGTTCTCGGGAATGAGTCTGATCTTCCACGTTGATGTTCTCCCGAGATC 1820
QY 1029 ATTTTGTAGTAACCTGTGAGACACAAATCAAGAGTTTGTGGAGTGAAGTGGTCAATG 1088
DB 1821 AGTATTTACGTGACTGTCCGGCCATAAGCAGGAAGTTTTCGGGACTCCGGTGGAAACCG 1880
QY 1089 ATAAACGTGAGTTGGATCTGGAGGAATGACAAATGTTTGTGGATCAACACT 1148
DB 1881 AAGATGGCCAACTAGCTTCAGGCGGAATGACAACTGATGTTGGTGGGACAGCTGA 1940
QY 1149 CAACCCAGGCTGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCAATGT 1208
DB 1941 ATGAGACACGCTTTATCGTTCTCGGACCATACCGCGCTGAGGCGCATCGCTGT 2000
QY 1209 CTCTCATCTTATGAGATCTTCTGATCTGGAGGAGAACTGACAGATAGATGTTTCTGTT 1268
DB 2001 CAGCCCATCAACCACTCTCTCGCTCAGGTGGAGGCGCGGCGGACCATCAAT 2060
QY 1269 TTTGGAATCAACCAAACTCACTACCTGATGATGACACTGGAAGTCAAGTTTGA 1328
DB 2061 TCTGGAACATGTAATGTTTCACTGATCAGGAAGTTGATACGGTACGAGGTGTGA 2120
QY 1329 ATCTTGTCTGTTCCAAATGTTCAAGAACTAGTAAGCAACATGAGGTTACTCCGAAACC 1388
DB 2121 ACCTGGCATGTTGGAAGTCTGAGGAACTCGGAGAAATCATGATACACAGGCTACAGTCAAAAC 2180
QY 1389 AGATTATTTGTTGGAGTACCCCACTATGTTCAAGCTGGGCTCTTACCGGCCATCTT 1448
DB 2181 AGATTATTTGTTGGAGTACCCCACTATGTTCAAGCTGGGCTCTTACCGGCCATCTT 2240
QY 1449 ATAGGGTCTCTATCTTTGCCATCTCTCCAGATGGACAGACTATTTGTAAGTGGAGTGGAG 1508
DB 2241 TCCGTGTTCTTACCTTGGATGAGCCCGCATGAGACCGTCTGTTACCGGAGCTGTGT 2300
```

```
QY 1509 ATGAAACGCTTAGGTTCTGGAATGTTTC 1537
DB 2301 ACGAACGTTGAGATTCGGAATTTTC 2329
```

## RESULT 11

ABT19494  
ID ABT19494 standard; DNA; 3993 BP.

XX ABT19494;

AC ABT19494;

DT 16-APR-2003 (first entry)

XX Aspergillus fumigatus essential gene #1852.

XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;

XX cancer; contamination; biofilm; antibody; immune response; ds.

XX Aspergillus fumigatus.

OS WO200286090-A2.

PN 31-OCT-2002.

PD 23-APR-2002; 2002WO-US13142.

XX 23-APR-2001; 2001US-285697P.

PR 27-APR-2001; 2001US-287066P.

PR 05-JUN-2001; 2001US-295890P.

PR 09-JUL-2001; 2001US-303899P.

PR 31-AUG-2001; 2001US-316362P.

XX (ELIT-) ELITRA PHARM INC.

XX Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;

PI WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of

XX Aspergillus fumigatus, useful for treating or preventing infections by

XX e.g. cancer -

XX Disclosure; Page -; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of  
XX essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
XX the invention are used to treat or prevent infections by a pathogenic  
XX organism such as A. fumigatus, to treat a non-infectious disease in a  
XX subject (e.g. cancer), to prevent or contain contamination of an object  
XX by A. fumigatus, or to prevent or inhibit formation on a surface of a  
XX biofilm comprising A. fumigatus. The polynucleotides are useful for  
XX expressing recombinant protein for characterisation, screening or  
XX therapeutic use, as markers for host tissues in which the pathogenic  
XX organisms invade or reside, for comparing with the DNA sequence of A.  
XX fumigatus to identify duplicated genes or paralogues having the same or  
XX similar biochemical activity and/or function, for comparing with DNA  
XX sequences of other related or distant pathogenic organisms to identify  
XX potential orthologous essential or virulence genes, for selecting and  
XX making oligonucleotides for attachment to a nucleic acid array for examination  
XX of expression patterns, for raising anti-protein antibodies, as an  
XX antigen to raise anti-DNA antibodies or to elicit another immune  
XX response, and for identifying polynucleotides encoding the other protein  
XX with which binding occurs or to identify inhibitors of the binding  
XX interaction. The polypeptides may be used to raise antibodies or to  
XX elicit immune response, as a reagent in assays designed to quantitatively  
XX determine levels of the protein in biological fluids, as a marker for  
XX host tissues in which pathogenic organism invade or reside, and to  
XX isolate correlative receptors or ligands in the case of virulence  
XX factors. This polynucleotide sequence represents one of the essential  
XX genes of Aspergillus fumigatus of the invention.



QY 1039 TAACTGTCAGACACAAATCAGAGGTTTGTGAGCTGAAGTGGTCAATGATGATACCGTGA 1098  
Db 12011 CAGATTGCCGACATCGCGAGGAGTGTGGGACTGAAATGGTCAACGGATATCAATA 12070  
QY 1099 GTTGCATCTGAGGAATGACAAATTTGTTTGGGAATCAACACTCAACCCAGCC 1158  
Db 12071 CTTGCCAGTGGCGGCAAGATATCGGTTGATGTGTGGATCAGCATTCGGTGAATCC 12130  
QY 1159 TGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGTCATGCTCTCATCT 1218  
Db 12131 CATAATCATACACAGGAGCATATGGCGCTGTAAAGCGATCGCGTGTGCGCGCATCA 12190  
QY 1219 TCATGATCTTGTGATCTGAGGAGGAACTGCAGATAGATGATTCGTTTTCGAATAC 1278  
Db 12191 CACGAGATCTTGGCAGCGCGGTGGAAACGGCGATAGGTGATCCGTTTCTGGAATAC 12250  
QY 1279 AACCAAACTCACACTTAGCTGATGAGACATGGAAGTCAAGTTTGCATCTGTGCTG 1338  
Db 12251 GCTGACGGGCCAGCCATCGAGTGGTGGACACGGGCTCGCAGGTTTGCATCTGGCGCTG 12310  
QY 1339 GTCCAAAATGTCACAGGACTAGTAAGCACACATCGGTACTCCAGAACCAAGATATTGT 1398  
Db 12311 GTCCAAAGCACTCTCGAGCTGGTCTCCAGCAGCGCTACTCGCAGAACCAAGATATTGT 12370  
QY 1399 TTGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCATACTTATAGGTTCT 1458  
Db 12371 GTGGAATATCCCTCCCTGACCAAGTGGCCAACTGACGGGCAATTCGTATCGTGTCT 12430  
QY 1459 CTATCTTGGCATCTCCAGATGGACAGACTATTGTAAGTGGAGCTGGAGTGAACGCT 1518  
Db 12431 CTATCTGGCGCTGAGTCCCGATGGTGAAGCTATTGTTACGGGCGCGGCGAGACGCT 12490  
QY 1519 TAGGTTCTGGAATGTTTTC 1537  
Db 12491 GCGATTTGGAAGCTATTC 12509

## RESULT 13

ABL18423  
ID ABL18423 standard; DNA; 1369 BP.

XX ABL18423;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 6742.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers BW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -

XX Claim 1; SEQ ID NO 6742; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABE57737-ABE72072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1369 BP; 304 A; 394 C; 394 G; 277 T; 0 other;

Query Match 12.4%; Score 248; DB 23; Length 1369;  
Best Local Similarity 55.9%; Pred. No. 1.5e-47;  
Matches 514; Conservative 0; Mismatches 400; Indels 6; Gaps 2;

QY 636 TTCTCGATCGCCTTATAAGTTTGGATGCACTGCTTTGCAAGATGATTTTATCTGA 695  
Db 396 TGCACGTGCGCCGTACAAGATCTCGATGCGCGGAGCTGCAAGACGACTTCTACCTGA 455  
QY 696 ATCTGATGATTTGCTTCCACCAATGTTTGGTCTTGGTAACTGTGCTATC 755  
Db 456 ACCTTATGACTGCTGCTGAAACACCTTGGCGCTTGGTCTCGGCTGCTCGTTATC 515  
QY 756 TCTGGAATGCTTTAGCAGCAAGTAACTAAATATGATTTGGGGTTGATGATG 815  
Db 516 TCTGGAGTGGCTGAGTGGTCAAGTCAAGCTATGCGACTTCAACAATAGGATAACC 575  
QY 816 TTTGTTCTGTTGGGCTTCAAGCT---GGTACTCATCTTCTGTTTGGAGTAACTA 872  
Db 576 TGATTACAGCAGTGTAGCTGGCAGCGGTGAAGCGCGGAGTGGCCATTTGGCACCAGTCCG 635  
QY 873 GTAAAGTTTCAAGTGGGATGCAAGATGCAAGAGATTAAGTCAATGGAGGGCCATC 932  
Db 636 GTTACGTGACCATTTGGGATGCGAGAACCAAAAGCAGATAAACCGACTGGAAGACACT 695  
QY 933 GGTTCATGTTGGGCTTGGCTGAGTTCATCTCTTTGTTCTTGGTGGAGCGGATA 992  
Db 696 CGCGCGTGTAAACGGCTAGCTTGGTGGGCAACCGATTTGGCAGTGGCTCCCGGACC 755  
QY 993 AGAATATTTATCAACGAGATATACGCA---CACAAAGAGATTTTGTAGTAACTGT 1049  
Db 756 GCTCAATTTGACGGGGATATCCGCAATCCGCGAGCAGACATACCGCTGCTGGCG 815  
QY 1050 GACAAATCAGAGTTTGGACTGAAGTGGTCAATATGATTAACGTTGAGTGGACTCTG 1109  
Db 816 GACAAAGTGAAGTGTGTGGCTTACAGTGGTGGCCAGCAATCGATCTTGGCCAGCG 875  
QY 1110 GAGGAATGACAAATTTGTTTGGATCAACACTCAACCCAGCCTGTCTCAAGT 1169  
Db 876 GTGACGAGCAACCGCTGTTGTTGGACCGAGTGGCGGAGCCCATTTACGCT 935  
QY 1170 ACTGTAGCACAACAGCAGCTGTTAAAGCTATTGATGAGTCTCTCATCTTCATGAGACTTC 1229  
Db 936 TCGACGAGCAACAGCGGTGTCAAAGCGCTGGGCTGTGCGCACATAAGTCCGGTCTTC 995  
QY 1230 TTGCATCTGGAGGAGGAACTGCAGATAGATGTTTGGTATTAACACACAACT 1289  
Db 996 TGGCCAGTGGCGGTGGGTCCCGATCGCTGCTTGGCTTTTGGACGCTCTCACCGGA 1055  
QY 1290 CACACCTTAGCTGTATGGACACTGGAAGTCAAGTTTGGATCTTGTGTGTCCTCAAAATG 1349  
Db 1056 AGCTGGTGAAGTGCATCAACACCGCGCCAGATCAGAACCTTGGCTTGGCGGAGGATT 1115  
QY 1350 TCAACGAACTAGTAAGCACACATGGTACTCCAGAACCAAGATTTGTTGGAGATACC 1409  
Db 1116 CCGAGAGCTGTTTACCAACCAACGAGCAGCGGAGCAGAGGTCATTTGCTTGGCGCTATC 1175  
QY 1410 CCAGTATGTCAAAGCTGGCGACTCTTATACCGGCCATATCTTATAGGTTTCTCTATCTGCCA 1469  
Db 1176 CTTGCTGAAGCAGATGGCCAGGCTGTCGCGGACACACAGCGGCTACTCCACCTGTCTCAG 1235

QY 1470 TCTCTCCAGATGACAGACTATTGTAACTGGAGCTGGAGATGAACACGCTTAGGTTCTGGA 1529  
 Db 1236 TGAGTCGGGATACGAGTCCATAGTGAATGGGCGCAGATGAGACGCTTCGCTTCTGGA 1295  
 QY 1530 ATGTTTCCCTTCCCTAAA 1549  
 Db 1296 CCGTTTTTACCAACAGAAA 1315

## RESULT 14

ABL18422  
 ID ABL18422 standard; DNA; 3369 BP.

XX ABL18422;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 6739.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Claim 1; SEQ ID NO 6739; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL1840-ABL16175), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pot\_sequences.

XX Sequence 3369 BP; 802 A; 899 C; 924 G; 744 T; 0 other;

XX Query Match 12.4%; Score 248; DB 23; Length 3369;

XX Best Local Similarity 55.9%; Pred. No. 2.1e-47;

XX Matches 514; Conservative 0; Mismatches 400; Indels 6; Gaps 2;

QY 636 TTCTCTGATCGCCCTTAAAGTTTGGATGACCTGCTTGGCAAGATGATTTTATCTGA 695

Db 1396 TGCCACGTCGCCCGCTACAGATATCTCGATGCGCGGAGCTGCAAGACGCTTCTACTGA 1455

QY 696 ATCTGGTAGATGTGCTTCAACAATGTGTTGGTGTGTTGGTAACTGTGTCTATC 755

Db 1456 ACCATTATGATGTGCTGCAAAACACCTGGCCGTGCTCGGGTGTCTCGTTATC 1515

QY 756 TCTGGAATGCTTTGAGCAGCAAGGATAAATATGTGATTTGGGGGTGATGATGTG 815

Db 1516 TCTGAGTGGCGTCAGTGTGAGTGACGCGCATGTGCGACTCAACAATGAGATAACC 1575  
 QY 816 TTTGTTCTGTGTGGTGGCTCAACGT---GGTACTCATCTGTGTGTGGAACTAAACAATG 872  
 Db 1576 TGATTACAGCAGTTAGCTGGCAGGTGAAGCGCGGAGTGGCCATTGGCACCAGTCCG 1635  
 QY 873 GTAAAGTTTCAGATTTGGGATGCGAGAGATGCAAGAGATAAGATCAATGAGGGCCATC 932  
 Db 1636 GGTAACGTGACCATTTGGGATGCGGAGAACCAAAAGCAGATAAACCGACTGGAAGACACT 1695  
 QY 933 GGTTACGTGTGGGCGCTTGGCCCTGGAGTTTCATCTCTTTTGTCTTCTGTGGTGA 992  
 Db 1696 CGGCGCTGTAAACGGCGCTAGCTTGGTGGCGCAACCGATTGGCCAGTGGCTCCGCGACC 1755  
 QY 993 AGAATATTATCAACGAGATATAGCA---CAAGAAGATTTTGTAGTAAACTGTCAG 1049  
 Db 1756 GCTCAATTTTGCAGCGGATATCCGCAATCCGCGAGCAGATACCCCGTGTGTGGCG 1815  
 QY 1050 GACAAATCAGAGTTTGTGGACTGAAAGTGGTCAATATGATAACCGTGGATGGCATCTG 1109  
 Db 1816 GACAAAGTAGAGTGTGTGGCTACAGTGGTGGCCGCAATCGATACTTGGCCAGCG 1875  
 QY 1110 GAGGAATGACAAACAATTTGTTTGGATCAACACTCAACCCAGCTGTCTCAAGT 1169  
 Db 1876 GTGGCAGCGCAACCGCTGTGTGGTGGACCGACGATTGGCCGAGCCCAITTTACGCCT 1935  
 QY 1170 ACTGTGAGCACACAGCAGCTGTTAAAGCTATTGATGCTCTCTCTCATCTTTCATGGA 1229  
 Db 1936 TCGACGAGCACACAGCGCTGTGTCAGGCGCTGGCTGGTCCGACATAGTCCGGTCTTC 1995  
 QY 1230 TTGCATCTGGAGGAGGAACTGCGAGATAGATGATTCGTTTTTGGATACAAACCAAACT 1289  
 Db 1996 TGGCCAGTGGCGTGGGTCCGCGATCGCTGTGGCTTTTGGAACTGCTCACCGGGA 2055  
 QY 1290 CACACCTTAGCTGTATGACACTGGAAGTCAAGTTTGCATCTTGTCTGTGTCAAAATG 1349  
 Db 2056 AGCTGGTGAAGTGTCAACACCGCGCCAGATCAGCACTGGCTGGCGCCAGGATT 2115  
 QY 1350 TCACGAACTAGTAAGCACATGGGTACTCCCGAAGACCAAGATTTGTTGGAGATACC 1409  
 Db 2116 CCCGAGAGCTGGTTACCAACCCACGACACGCGGAGCCACAGGTCAATTGCGCGCTATC 2175  
 QY 1410 CCACATGTCAAAAGCTGGCGACTCTTACCGGCCATCTTATAGGGTCTCTATCTTGCCA 1469  
 Db 2176 CTTGCTGAACGATGCGCAGGCTGTCCGGGACACACAGCGGTTACTCCACTGTGAG 2235  
 QY 1470 TCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAGATGAACCGCTTAGTGTGGA 1529  
 Db 2236 TGAGTCCGGATAACGAGTCCATAGTGAAGTGGCGCGCAGATGAGACGCTTGGCTTCTGGA 2295  
 QY 1530 ATGTTTCCCTTCCCTAAA 1549  
 Db 2296 CCGTTTTTACCAACAGAAA 2315

## RESULT 15

ABL05448

ID ABL05448 standard; cDNA; 5058 BP.

XX AC ABL05448;

XX AC ABL05448;

XX DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 10826.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN

XX







QY 1037 AGTAACTGTGAGGACAAATCAGAGGTTTGTGGACTGAAGTGGTCTATATGATACCCGT 1096  
Db 1099 GCCACATGAGTGGCCATAGCCAGGAGTATGTGGCTGGCTGGCCCGCCAGATGGACGA 1158  
QY 1097 GAGTTGGCATCTGGAGGAATGACAAATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1144  
Db 1159 CATCTGCAAGCGGTGGCAATGATAATTTGTCAAGTGTGGCTGGTGGTGGTGGTGGTGG 1218  
QY 1145 CACTCAACCCAGCGCTGCTCAAGTACTGTGAGCAGACAGCAGCTGTTAAAGCTATTGCA 1204  
Db 1219 AGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1278  
QY 1205 TGGTCTCTCATCTTCTGAGCTTCTTGGATCTTCTGAGGAGGAACTGCGAGATAGATATT 1264  
Db 1279 TGGTCTCTCTGAGCTTCTTGGATCTTCTGAGGAGGAACTGCGAGATAGATATT 1338  
QY 1265 CGTTTTGGATACAAACCAAACTCACCTTCTGAGTGTGATGACAGCTGGAAGCTCAGGTT 1324  
Db 1339 CGCAATTTGGAGCTGCTCTGGAGCTGCTGAGTGTGATGAGTGTGATGAGTGTGATGAG 1398  
QY 1325 TGCAATCTTGTCTGGTCCAAATGTCACCAAACTAGTAAAGCAGACATGGGTACTCCAG 1384  
Db 1399 TGTCTCATCTCTGCTCTCCCACTATAAGGAGCTCATCTCAGGCCATGGCTTTGCCAG 1458  
QY 1385 AACAGATATTGTTGGAGTACCCCACTATGTCAAAGCTGGGAGCTCTTACCGCCAT 1444  
Db 1459 AACAGCTGGTATTGGAGTACCCCACTATGTCAAAGCTGGGAGCTCTTACCGCCAT 1518  
QY 1445 ACTTATAGGTTCTCTATCTTGGCTCTCTCCAGATGGAGAGCTATTGTAAGTGGAGCT 1504  
Db 1519 ACAGCCGGTCTCTGAGTCTCACCATGATGATCCAGAGCGGGCCAGTGGCATCTGCAGCA 1578  
QY 1505 GGAGATGAAGCTTAGTCTGG 1528  
Db 1579 GCCGATGAGACTCTGGGCTCTGG 1602

## RESULT 2

PCT-US95-01806-1  
; Sequence 1, Application PC/TUS9501806  
; GENERAL INFORMATION:  
; APPLICANT: Weinstein, Jasminder  
; TITLE OF INVENTION: Novel Mammalian Cell Cycle Protein  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Inc./Patent Operations/RBW  
; STREET: 1840 Dehavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01806  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1767 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 193..1692  
PCT-US95-01806-1

Query Match

10.0%; Score 200.4; DB 5; Length 1767;

Best Local Similarity 52.8%; Pred. No. 9.7e-42;  
Matches 488; Conservative 0; Mismatches 421; Indels 15; Gaps 2;  
QY 620 AAGGCTCTAGGAAGTTCCTCGATCGCCTTATAAGGTTTGGATGACACCTGCTTTGCAA 679  
Db 679 AAGGCTTCAGATACATTCCTTCCCTGCCAGAGGATTTCTTATGATGCCCTGAAATCCGG 738  
QY 680 GATGATTTTATCGAATCTGGTAGATGTTGTTTCAACAAATGTTGGTCTGTTGGTTG 739  
Db 739 AATGACTACTACCTGAATCTTGTGATTTGGAGCTCTCGAAATGTTATGCTGTGGCACTG 798  
QY 740 GGTAACTGTGCTATCTCTGAATGCTTGTAGCAGCAAGGTAACTAAATTTATGTTGATTG 799  
Db 799 GACAACTGTGTTACTTATGGAACGCTGTTCCGCTGACATCTCTGAGCTGTTGCAATG 858  
QY 800 G---GGGTTGATGATGTTGTTGTTCTGTTGGTCTCAAAGTGTGCTCTCATCTTGTCT 856  
Db 859 GAGCAGCTCTGGGACTACATATCATCTCGTGGCTCGATCAAAGAGGGCACTACTCGCT 918  
QY 857 GTTGAACCTAAATGTTAAAGTTCAGATTTGGATTCAGCAGATGCAAGATCAAGATAAGA 916  
Db 919 GTGGGACACGATATGCTGAGTGCAGCTATGGATGTGAGCAGCAGAGAACGCTTCGA 978  
QY 917 TCAATGAGGGCCATCGTTTACGTGTGGGCTTGGCTGGAGTTCACTCTCTTTGTCT 976  
Db 979 AACATGACAGCCACTCTGCTCGAGTAGCTCCCTGAGTTGGAACAGCTATATCTGTCA 1038  
QY 977 TCTGTGAGCGGATAAGAAATTTATCAAGAGATATACGACACAAAGAAAGATTTTCTT 1036  
Db 1039 AGTGGTTTCAAGATCTGGCCATCCACCACCATGTTTCAGTAGCAGAAACACCATGTG 1098  
QY 1037 AGTAACTGTGAGGACAAATCAGAGGTTTGTGGACTGTGAAGTGGTCTATATGATAACCGT 1096  
Db 1099 GCCACACTGAGTGGCCATAGCCAGGAAGTATGTGGCTGCGCTGGGCCCCAGATGAGCA 1158  
QY 1097 GAGTTGGGATCTGGAGGAATGACAAATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 1144  
Db 1159 CATCTGGCAAGCGGTGGCAATGATAACTGTCAACGTTGGCTTAGTGGTCTCTGGAGAA 1218  
QY 1145 CACTCAACCCAGCTGTCTCAAAGTACTGTGAGCAGACACAGCAGCTGTTTAAAGCTATTGCA 1204  
Db 1219 AGTGGCTGGTTCCTCTGAGACATTTCACTCAACATCAAGTGTCTGTCAGGCTGTTGCA 1278  
QY 1205 TGGTCTCTCATCTTCTGAGCTTCTTGCATCTGGAGGAGGAACCTGAGATAGATATT 1264  
Db 1279 TGGTCTCTCTGAGCTTCTTGCATATCTGCAACAGGAGGAGTACAGTGAACGACATTT 1338  
QY 1265 CGTTTTTGGATACAAACCAAACTCAACCTTACTGTATGGACACTGGAAGTCAAGTTT 1324  
Db 1339 CGCAATTTGGAACGCTGCTCTCTGGAGCTGCTGAGTGTCTGATGTGCTATTCAGGTTG 1398  
QY 1325 TGCAATCTTGTCTGTTCCAAATGTCACGAACTAGTAAAGCAGACATCTGGGTACTCCAG 1384  
Db 1399 TGTCTCATCTCTGCTCTCCCACTATAAGGAGCTCATCTCAGGCCATGGCTTTGCCAG 1458  
QY 1385 AACGATATTGTTTGGAGTACCCCACTATGTCAAAGCTGGGAGCTTCTTACCGCCAT 1444  
Db 1459 AACGAGCTGGTATTGAGTACCCCACTATGGCCAGGTTGGCAGAGCTCAAAGTCAAC 1518  
QY 1445 ACTTATAGGTTCTCTATCTTGGCTCTCTCCAGATGGAGAGCTATTGTTAACTGGAGCT 1504  
Db 1519 ACAGCCGGTCTCTGAGTCTCACCATGATGATCCAGAGCGGGCCAGTGGCATCTGCAGCA 1578  
QY 1505 GGAGATGAAGCTTAGTCTGG 1528  
Db 1579 GCCGATGAGACTCTGGGCTCTGG 1602

## RESULT 3

US-08-988-856B-3  
; Sequence 3, Application US/08988856B  
; Patent No. 6291642  
; GENERAL INFORMATION:

APPLICANT: Weinstein, Jasminder  
TITLE OF INVENTION: No. 6291642el Mammalian Cell Cycle Protein  
FILE REFERENCE: 06843.0026-04; A-283 D  
CURRENT APPLICATION NUMBER: US/08/988,856B  
CURRENT FILING DATE: 1997-12-11  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1700  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-988-856B-3

Query Match 9.9%; Score 198.8; DB 3; Length 1700;  
Best Local Similarity 52.7%; Pred. No. 2.4e-41;  
Matches 487; Conservative 0; Mismatches 422; Indels 15; Gaps 2;

QY 620 AAGGCTCCAGAGGTTCTCGATCGCTTAAAGTTTGGATGCACTGCTTTGCAA 679  
DB 603 AAGACCTGCGTTACATTCCTCCCTGCCAGACCGGTATCTCGATGCGCTGAATCCGA 662  
QY 680 GATGATTTTATCTGAATCTGTAGATGCTCTTACACAAATGTTGGCTCTTGGTTG 739  
DB 663 AATGACTATTAACCTTGTGATGAGTTCTGGGAATGTAAGTGGCTGCACTG 722  
QY 740 GGTAACTGTCTATCTCTGGAATGCTTGTACGACCAAGTAACATAATATGTAATTG 799  
DB 723 GACACAGTGTGTACCTGTGGAGTGAAGCTCTGTGTGACATCTCTGAGCTTTTGCAAATG 782  
QY 800 G---GGGTTGATGATGTTGTTCTGTTGTTGGCTCAACGCTGTAATCTCTGCT 856  
DB 783 GAGCAGCTGCGGGAATATATATCTCTGTGCGCTTGGATCAAGAGGCGCACTACTTGCT 842  
QY 857 GTTGGAACTAACAAATGTTAAAGTTTCAAGTTTGGATGCAAGATGCAAGAGTAAGA 916  
DB 843 GTGGGACACGAGTGTGAGGTGCACTATGCTGTGGCTTGAATCAAGAGGCGCACTACTTGCT 902  
QY 917 TCAATGAGGGCCATCGTTACGTGTGCGGCTTGGCTGAGTTCACTCTTTTGTCT 976  
DB 903 AATATGACCAAGTCACTCTGCGGAGTGGCTCCCTAAGCTGGAACAGCTATATCTCTGCC 962  
QY 977 TCTGGTGGAGCGGATAAGATAATTTATCAACGAGATATACGCAACAAGAGATTTTGT 1036  
DB 963 AGTGGTTTCAAGTTCTGGCCACATCCACCACCATGATGTTGGGTAGCAGAACCATGTG 1022  
QY 1037 AGTAACTGTGAGACACAAATCAGAGGTTTGTGATGTAAGTGGTCAATATGATACCT 1096  
DB 1023 GCCACACTGAGTGGCCACAGCCAGGAAGTGTGTGGCTGCGCTGGGCCCCCAGATGCA 1082  
QY 1097 GAGTTGGCATCTGGAGGAATGACAAATAATGTTGTTTGGTAATCAACACTCAACCCAG 1156  
DB 1083 CATTTGGCCAGTGTGTGATGATTAATCTTGTGCTCAATGTGTGGCTAGTCTCTGGAG 1142  
QY 1157 -----CCTGTCTCAAGTACTGTGAGCAGACAGCTGTTAAAGCTATTGCA 1204  
DB 1143 GGTGGCTGGGTTCTCTGACAGATTCACCCAGCATCAAGGGCTGTCAAGGCGGTAGCA 1202  
QY 1205 TGGTCTCTCATCTTCAATGACTTCTGTGATCTGAGGAGGAACTGCAAGATAGATATT 1264  
DB 1203 TGGTGTCTCTGGAGTCCAAATGCTCTGGCAACAGAGGGGGCACTGATGACACATT 1262  
QY 1265 CGTTTTTGGAAATACAAACCAAACTCACACCTTAGCTGTATGACACTGGAAGTCAAGTT 1324  
DB 1263 CGCATCTGGAATGTGCTCTGCGGCTGTCTGAGTGGCTGATGCCATCCAGGTG 1322  
QY 1325 TGAATCTTGTCTGTCGTCGCAAAATGTCAGGACTAGTAGCAGACATGGTACTCCAG 1384  
DB 1323 TGCTCCATCTCTGCTGCTCCCATTAACAGAGCTCATCTCAGGCCATGCTTGCACAG 1382  
QY 1385 AACAGATTTATTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCAT 1444  
DB 1383 AACAGCTAGTTATTTGGAGTACCCCAACCATGTCGCAAGGTGGCTGAATCAAGAGTCA 1442

QY 1445 ACTTATAGGTTCTCTATCTTCCATCTCTCCAGATGACAGACTATTGTAACTGAGCT 1504  
DB 1443 ACATCCGGTCTGTAGTCTGACCATGAGCCAGATGGGCGCCACAGTGGCATCCGAGCA 1502  
QY 1505 GGAGATGAACAGCTTATAGTTCTGG 1528  
DB 1503 GCAGATGAGACCTGAGGCTATGG 1526

RESULT 4  
PCT-US95-01806-3  
Sequence 3, Application PC/TUS9501806  
GENERAL INFORMATION:  
APPLICANT: Weinstein, Jasminder  
TITLE OF INVENTION: Novel Mammalian Cell Cycle Protein  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Angen Inc./Patent Operations/RBW  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01806  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 117..1616  
PCT-US95-01806-3

Query Match 9.9%; Score 198.8; DB 5; Length 1700;  
Best Local Similarity 52.7%; Pred. No. 2.4e-41;  
Matches 487; Conservative 0; Mismatches 422; Indels 15; Gaps 2;

QY 620 AAGGCTCCAGAGGTTCTCGATCGCTTAAAGTTTGGATGCACTGCTTTGCAA 679  
DB 603 AAGACCTGCGCTTACATTCCTTCCCTGCCAGACCGTATCTCGATGCGCTGAATCCGA 662  
QY 680 GATGATTTTATCTGAATCTGTAGATGTTCTTACACAAATGTTGGCTGTTGGTTG 739  
DB 663 AATGACTATTAACCTTGTGATGAGTTCTGGGAATGTAAGTGGCTGCACTG 722  
QY 740 GGTAACTGTCTATCTCTGGAATGCTTGTAGCAGCAAGTAACATAATATGTAATTG 799  
DB 723 GACACAGTGTGTACCTGTGGAGTGAAGCTCTCTGTGACATCTCTGAGCTTTTGCAAATG 782  
QY 800 G---GGGTTGATGATGTTGTTGTTGTTGGCTCAACGCTGTAATCTCTGCT 856  
DB 783 GAGCAGCTGCGGGAATATATATCTCTGTGCGCTTGGATCAAGAGGCGCACTACTTGCT 842  
QY 857 GTTGGAACTAACAAATGTTAAAGTTTCAAGTTTGGATGCAAGATGCAAGAGTAAGA 916  
DB 843 GTGGGACACGAGTGTGAGGTGCACTATGCTGTGGCTTGAATCAAGAGGCGCACTACTTGCT 902  
QY 917 TCAATGAGGGCCATCGTTACGTGTGCGGCTTGGCTGAGTTCACTCTTTTGTCT 976  
DB 903 AATATGACCAAGTCACTCTGCGGAGTGGCTCCCTAAGCTGGAACAGCTATATCTCTGCC 962  
QY 977 TCTGGTGGAGCGGATAAGATAATTTATCAACGAGATATACGCAACAAGAGATTTTGT 1036



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,675  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.45501  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BMB UT  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4164 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: PAT-1  
US-08-204-675-1

Query Match 2.3%; Score 46.4; DB 1; Length 4164;  
Best Local Similarity 53.3%; Pred. No. 0.035;  
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATTCGAAAGATGATACAAATGGTAAATTTATTGCTTGGACTTATACAT 1881  
DB 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTCTGCTAATAATATATAT 3900

QY 1882 GCATTGATGAGTGTAGCCAAAGTTTTTTTATTACTTTTTTTCTTTCTTTTGGAT 1941  
DB 3901 GTAAAGTACGCTTTTGTGGAATTTTTTAAACCTTTGTTATTTTTTTCTTCATCC 3960

QY 1942 AGTGCTCTCGCTGATTTATATATTTTAAAGTGGTTAAACAGAGAAAAAATAA 2001  
DB 3961 GTAACTCTTACCTTTCTTTTATTACTTTCTTAAATCCAAATACAAACATATAAATAA 4020

QY 2002 AAAA 2005  
DB 4021 TAAA 4024

RESULT 8  
US-08-660-754-1  
Sequence 1, Application US/08660754  
Patent No. 5843772  
GENERAL INFORMATION:  
APPLICANT: Devine, Scott E.  
APPLICANT: Boeke, Jef D.  
APPLICANT: Braiterman, Lelita T.  
TITLE OF INVENTION: In Vitro Transposition of Artificial  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,754  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/204,675  
FILING DATE: 02-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.45501  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BMB UT  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4164 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: PAT-1  
US-08-660-754-1

Query Match 2.3%; Score 46.4; DB 2; Length 4164;  
Best Local Similarity 53.3%; Pred. No. 0.035;  
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATTCGAAAGATGATACAAATGGTAAATTTATTGCTTGGACTTATACAT 1881  
DB 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTCTGCTAATAATATATAT 3900

QY 1882 GCATTGATGAGTGTAGCCAAAGTTTTTTTATTACTTTTTTTCTTTCTTTTGGAT 1941  
DB 3901 GTAAAGTACGCTTTTGTGGAATTTTTTAAACCTTTGTTATTTTTTTCTTCATCC 3960

QY 1942 AGTGCTCTCGCTGATTTATATATTTTAAAGTGGTTAAACAGAGAAAAAATAA 2001  
DB 3961 GTAACTCTTACCTTTCTTTTATTACTTTCTTAAATCCAAATACAAACATATAAATAA 4020

QY 2002 AAAA 2005  
DB 4021 TAAA 4024

RESULT 9  
US-08-796-364-1  
Sequence 1, Application US/08796364  
Patent No. 5968785  
GENERAL INFORMATION:  
APPLICANT: Devine, Scott E.  
APPLICANT: Boeke, Jef D.  
APPLICANT: Braiterman, Lelita T.  
TITLE OF INVENTION: In Vitro Transposition of Artificial  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/796,364  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/204,675  
 FILING DATE: 02-MAR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kagan, Sarah A.  
 REGISTRATION NUMBER: 32,141  
 REFERENCE/DOCKET NUMBER: 01107.45501  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202.508.9100  
 TELEFAX: 202.508.9299  
 TELEX: 197430 BMB UT  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4164 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 IMMEDIATE SOURCE:  
 CLONE: PAT-1  
 PCT-US95-02520-1

Query Match 2.3%; Score 46.4; DB 2; Length 4164;  
 Best Local Similarity 53.3%; Pred. No. 0.035;  
 Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
 QY 1822 CTCATTTTATATGAAAGATGATAACAAATGGGTAATTTATTTCTGTGGACTTATACAT 1881  
 DB 3841 CTTATCGTTAATCGAATGATATCTATTTAATCTGCTTTTCTGTCTAATAATATATAT 3900  
 QY 1882 GCATTGAGGAGTGTAGCCAAAGTTTTTATTTACTCTTTTCTTTCTTTCTTTTTCAT 1941  
 DB 3901 GTAAAGTAGCGCTTTTGTGTGAATTTTAAACCTTTGTTTATTTTCTTCATTC 3960  
 QY 1942 AGTGTCTCTCCGCAATTTTATATAATTTTAAGATGGGTTAACAGAGAAAAAATAAAA 2001  
 DB 3961 GTAACCTCTTACCTCTTTTACTTTTACTTTTCTTAAATCCAAATACAAATATAA 4020  
 QY 2002 AAAA 2005  
 DB 4021 TAAA 4024

RESULT 10  
 PCT-US95-02520-1  
 Sequence 1, Application PC/TUS9502520  
 GENERAL INFORMATION:  
 APPLICANT: The Johns Hopkins University  
 TITLE OF INVENTION: In Vitro Transposition of Artificial  
 TITLE OF INVENTION: Transposons  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: Banner, Birch, McKie, and Beckett  
 STREET: 1001 G Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20001  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/02520  
 FILING DATE: 02-MAR-1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.  
 REGISTRATION NUMBER: 32,141  
 REFERENCE/DOCKET NUMBER: 01107.49245  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202.508.9100  
 TELEFAX: 202.508.9299  
 TELEX: 197430 BMB UT  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4164 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 IMMEDIATE SOURCE:  
 CLONE: PAT-1  
 PCT-US95-02520-1

Query Match 2.3%; Score 46.4; DB 5; Length 4164;  
 Best Local Similarity 53.3%; Pred. No. 0.035;  
 Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
 QY 1822 CTCATTTTATATGAAAGATGATAACAAATGGGTAATTTATTTCTGTGGACTTATACAT 1881  
 DB 3841 CTTATCGTTAATCGAATGATATCTATTTAATCTGCTTTTCTGTCTAATAATATATAT 3900  
 QY 1882 GCATTGAGGAGTGTAGCCAAAGTTTTTATTTACTCTTTTCTTTCTTTCTTTTTCAT 1941  
 DB 3901 GTAAAGTAGCGCTTTTGTGTGAATTTTAAACCTTTGTTTATTTTCTTCATTC 3960  
 QY 1942 AGTGTCTCTCCGCAATTTTATATAATTTTAAGATGGGTTAACAGAGAAAAAATAAAA 2001  
 DB 3961 GTAACCTCTTACCTCTTTTACTTTTACTTTTCTTAAATCCAAATACAAATATAA 4020  
 QY 2002 AAAA 2005  
 DB 4021 TAAA 4024

RESULT 11  
 US-08-204-675-2  
 Sequence 2, Application US/08204675  
 Patent No. 5671170  
 GENERAL INFORMATION:  
 APPLICANT: Devine, Scott E.  
 APPLICANT: Boeke, Jef D.  
 APPLICANT: Braiterman, Lelita T.  
 TITLE OF INVENTION: In Vitro Transposition of Artificial  
 TITLE OF INVENTION: Transposons  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: Banner, Birch, McKie, and Beckett  
 STREET: 1001 G Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20001  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/204,675  
 FILING DATE: 02-MAR-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kagan, Sarah A.  
 REGISTRATION NUMBER: 32,141  
 REFERENCE/DOCKET NUMBER: 01107.45501  
 TELECOMMUNICATION INFORMATION:





TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4933 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: PAT-2  
US-08-796-364-2

Query Match 2.3%; Score 46.4; DB 2; Length 4933;  
Best Local Similarity 53.3%; Pred. No. 0.037;  
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881  
Db 4610 CTTATCGTTAATCGAATGTAATCTATTTAATCTGCTTTTCTTGCTAAATAATATAT 4669

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941  
Db 4670 GTAAAGTACGCTTTTGTGGAATTTTAAACCTTGTTATTTTCTTCTTCATCC 4729

QY 1942 AGTGCTCTCGCATTTATTTATATAATTTAAGATGGTTAAGACAGAAAAA 2001  
Db 4730 GTAACTCTCTACCTCTTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4789

QY 2002 AAAA 2005  
Db 4790 TAAA 4793

RESULT 14  
PCT-US95-02520-2  
Sequence 2, Application PC/TUS9502520  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University  
TITLE OF INVENTION: In Vitro Transposition of Artificial  
TITLE OF INVENTION: Transposons  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie, and Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02520  
FILING DATE: 02-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.49245  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202.508.9100  
TELEFAX: 202.508.9299  
TELEX: 197430 BBMB UT  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4933 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: PAT-2  
PCT-US95-02520-2

Query Match 2.3%; Score 46.4; DB 5; Length 4933;  
Best Local Similarity 53.3%; Pred. No. 0.037;  
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881  
Db 4610 CTTATCGTTAATCGAATGTAATCTATTTAATCTGCTTTTCTTGCTAAATAATATAT 4669

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941  
Db 4670 GTAAAGTACGCTTTTGTGGAATTTTAAACCTTGTTATTTTCTTCTTCATCC 4729

QY 1942 AGTGCTCTCGCATTTATTTATATAATTTAAGATGGTTAAGACAGAAAAA 2001  
Db 4730 GTAACTCTCTACCTCTTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4789

QY 2002 AAAA 2005  
Db 4790 TAAA 4793

RESULT 15  
US-09-890-848-16  
Sequence 16, Application US/09890848  
Patent No. 6573053  
GENERAL INFORMATION:  
APPLICANT: Firth, Greg  
APPLICANT: Odedra, Rajesh Muru  
TITLE OF INVENTION: Analysis Method  
FILE REFERENCE: PA9903  
CURRENT APPLICATION NUMBER: US/09/890,848  
CURRENT FILING DATE: 2001-08-03  
PRIOR APPLICATION NUMBER: PCT/GB00/00346  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: EP 99300873.9  
PRIOR FILING DATE: 1999-02-05  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 16  
LENGTH: 5277  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: plasmid  
US-09-890-848-16

Query Match 2.3%; Score 46.4; DB 4; Length 5277;  
Best Local Similarity 53.3%; Pred. No. 0.038;  
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881  
Db 4954 CTTATCGTTAATCGAATGTAATCTATTTAATCTGCTTTTCTTGCTAAATAATATAT 5013

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941  
Db 5014 GTAAAGTACGCTTTTGTGGAATTTTAAACCTTGTTATTTTCTTCTTCATCC 5073

QY 1942 AGTGCTCTCGCATTTATTTATATAATTTAAGATGGTTAAGACAGAAAAA 2001  
Db 5074 GTAACTCTCTACCTCTTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5133

QY 2002 AAAA 2005  
Db 5134 TAAA 5137

us-09-701-572-1.rni

Sun Jan 25 17:25:16 2004

Search completed: January 23, 2004, 21:08:23  
Job time : 142 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	326.2	16.3	1143	12	US-10-369-493-27884	Sequence 27884, A
2	303.8	15.1	1794	13	US-10-032-585-6595	Sequence 6595, Ap
3	277	13.8	1312	12	US-10-369-493-26005	Sequence 26005, A
4	277	13.8	1312	12	US-10-369-493-26011	Sequence 26011, A
5	277	13.8	1312	12	US-10-369-493-26013	Sequence 26013, A
6	264.6	13.2	1116	12	US-10-369-493-36526	Sequence 36526, A
7	263.8	13.2	1398	15	US-10-138-714-1038	Sequence 1038, Ap
8	263.8	13.2	1398	15	US-10-138-714-2038	Sequence 2038, Ap
9	263.8	13.2	1875	15	US-10-138-714-7038	Sequence 7038, Ap
10	263.8	13.2	1993	15	US-10-138-714-6038	Sequence 6038, Ap
11	263.8	13.2	3398	15	US-10-138-714-38	Sequence 38, Appl
12	263.8	13.2	3993	15	US-10-138-714-5038	Sequence 5038, Ap
13	225.6	11.2	1701	12	US-10-369-493-45642	Sequence 45642, A
14	213.6	10.6	1344	10	US-09-938-842A-988	Sequence 988, App
15	213.6	10.6	1344	12	US-09-938-842A-988	Sequence 988, App

Db 367 CAACAAGTAGAGTCAACAAGCTGTGACGTTGGAG---GACGACACGGTGGCCACGCTA 423  
Qy 827 GGTGGGCTCAACGCTGGTACTCATCTTGTGTTGGAACTAAACAATGGTAAAGTTCAAGTT 886  
Db 424 TCTTGGATACAAAGGGGACACACCTTGTCTATTGGTACCCATTAAGGACCTGTGTGACATC 483  
Qy 887 TGGGATGACGACAGATGCAAGAGATGAAGATCAATGAGGGGCCATCGTTAGCTGTCGGG 946  
Db 484 TGGGATGCTGAGAAAGCAAGAGCTTAAGAACGATGACTGCCCCATCTACTGCGAGGTTGGA 543  
Qy 947 GCCTTGGCCCTGGAGTTTCATCTCTTTTGTCTTCTGTGGACGGGATAGAAATATTTATCAA 1006  
Db 544 GCATTGCACTGGAACACTCATATTCTCACTGAGGATCCCGGATCCGCTTAATATACCAT 603  
Qy 1007 CGAGATATACGACACAGAGAGATTTTGTAGTAACTGTGACGACACAAATCAGAGTT 1066  
Db 604 CGGACCTCGAGGCGCAGATCATGTGCTCAAAAAGCTGTGTTGTCACAGCAGGAGTG 563  
Qy 1067 TGTGGACTGAAGTGGTCAATATGATAACCGTGAAGTGGCACTCTGGAGGAAATGACAA 1126  
Db 664 TGTGCTCTCAATGGAATTGCGAGATGGCCAACTTGGCCAGTGGAGGAAACGACAATAAG 723  
Qy 1127 TTGTTTGTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTGAGCACACAGCA 1186  
Db 724 TTGATGCTTGGGATAGCTTTCCGACACACCTCTATGGAAGTATTCCGGTCAACAGCA 783  
Qy 1187 GCTGTTAAAGTATTGCAATGCTCTCTCATCTTTCATGGAATCTTCTGTCATCTGGAGGAGGA 1246  
Db 784 GCCGTCAAGGATAGCTGGTCAACACATCAAGCGGTCTCTCGGCTCCGAGGCGGT 843  
Qy 1247 ACTGACAGATGATGTTTGGTTTGGATACAAACCAACACTCAACCTCAACCTTACTGTATG 1306  
Db 844 ACTGCTGACAGACGATCATCTTTTCCAGCACACCGTTCCGCGCACGCTACTCAACAGGTTG 903  
Qy 1307 GACACTGGAAGTCAAGGTTTGGAACTTGTCTGTGTCGCAAAAATGTCACGAACTAGTAAGC 1366  
Db 904 GACACGGGCTGAGGCTGTGAACTTGCCTGGTCCAGAACTCCAAACGAAATGTTCTCG 963  
Qy 1367 ACATAGGATGACTCCGAGAACAGATTAATGTTTGGAGATACCCCACTATGTCAAAGCTG 1426  
Db 964 ACGATGATATAGCCAGAACAGATAGTATTTTGGAAAGTACCCCTCCATGACGCAAGTC 1023  
Qy 1427 GCGACTCTACGGCCATCTATATAGGTTCTCTATCTTGGCATCTCTCCAGATGGACAG 1486  
Db 1024 GCCAGCTGACCGGTCAACATTCGCGTGTGTGTTGCTTGGCATGAGCCCGGACGCAAA 1083  
Qy 1487 ACTATTGTAAGTGGAGCTGGAGATCAAAACCGCTTAGGTTCTGGAATGTTTC 1537  
Db 1084 ACCGTGGTGACGGGTGCGGAGATGAGACATTGGGGTTCTGGAATCTCTTC 1134

RESULT 2  
US-10-032-585-6595  
; Sequence 6595, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6595  
; LENGTH: 1794  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-10-032-585-6595

Query Match 15.1%; Score 303.8; DB 13; Length 1794;

Best Local Similarity 58.1%; Pred. No. 3.6e-66;  
Matches 536; Conservative 0; Mismatches 387; Indels 0; Gaps 0;  
Qy 614 CCGGTTAAGGCTCTTAGAAGGTTCTCGATCGCCTTATAAGGTTTGGATGACCTGCT 673  
Db 805 CCGCAGAGAAGCCCTCGAACCAATATCAAGGTTCCATATAGAGTATTGGATGCAACAGAA 864  
Qy 674 TTGCAAGATGATTTTATCTGAATCTGATAGTCTTTCACACAATGTTGGCTGTT 733  
Db 865 CTATCTGATGATTTTATTTGAATTTAGTTGATGGGTCACACAGGATGACTAGCCGTA 924  
Qy 734 GGTTTGGTAACTGTGCTATCTCTGGAATGCTTTGAGCAGCAAGGTAACATAATTATGT 793  
Db 925 GGAATGGGCGACAGTGTGTATTATTTATGGGATGGCCCAACAATCAGTGGACAGATTTGT 984  
Qy 794 GATTTGGGGTGTGATGATTTGTGTTCTGTTGGGCTCAACGTTGGTACTCATCTT 853  
Db 985 AATTTCACAAACAGGACAAAGTCAAGTTTGAATGGATAGTACTGGTACCCATTG 1044  
Qy 854 GCTGTTGGAACATAAATGTTAAAGTTCAGATTTGGGATGACGCAAGATGCAAGAGATA 913  
Db 1045 GCAATAGGAAACATCAAGGGTTTGGTGGAGATATGGGATGCAACTCGAATCAAGTGCAATC 1104  
Qy 914 AGATCAATGGAGGGCCATCGGTTACGTTGCGGGCTTGGCTCGAGTTCATCTCTTTTG 973  
Db 1105 AGGACATGACAGGACACTCTCTACGGTTAGTTCAATAGCGTGAATGAACATATTA 1164  
Qy 974 TCTTCTGTTGACCGGATTAAGAAATTTATCAACGAGATATACGACACAAGAGATTTT 1033  
Db 1165 AGCAGTGGGAGTCGCGATCGAACCAATTTAAATCGCATGTGAGAAATTTGAAGATCACCTT 1224  
Qy 1034 GTTACTAACTGTCAAGGACACAAATCAAGGTTTGTGACCTGAAGTGTGATATGATTAAC 1093  
Db 1225 GTTAAATAATTTGATTAATCATAGCAAGAGTGTGACCTCAAGTGAATGTGGAAGAA 1284  
Qy 1094 CGTGAAGTTGGCATCTGGAGGAANTGACAAACAAATTTGTTTGTGGAATCAACACTCAACC 1153  
Db 1285 AACAACTAGCTAGTGTGGTGGCAACGATAATAATTTGTTGATGGGATGGGCTCAATCCT 1344  
Qy 1154 GAGCCTGCTCCTCAAGTACTGTGACGACACAGCAGCTGTTAAAGCTATTGCAATGCTCCT 1213  
Db 1345 AAACGTTGTCATCAATTTACAGACCAACTGCTGCCGTTAAAGCTATTGCTGTCGCG 1404  
Qy 1214 CATCTTCATGAGTCTTTCGATCTGGAGGAGAACTGCAGATAGATGATTTGTTTGG 1273  
Db 1405 CATCAACGGGGATTTTGGCATCGGTCGCGAACAGCCGATAGACTATAAAGACATGG 1464  
Qy 1274 AATACAAACCAAACTCAACCTTAGCTGTATGGACACTGGAAGTCAAGTTTGCATCTT 1333  
Db 1465 AATAAGTTGACTGGGAATTTAGTACATGATGTCAATCTGCGTCAAGTGTGTAATTTA 1524  
Qy 1334 GTCTGCTCCAAAATGTCAACGAACTAGTAAGCACAATGGGTACTCCAGAACCAAGATT 1393  
Db 1525 ATCTGCTCAAAAATTTCCACGAAATGGTTTCAACCCATGGTTATTCAAGAAATCAAAT 1584  
Qy 1394 ATTGTTGGAGATACCCCACTATGTCAAAGCTGGGACCTTTACCGGCCATCTTATAGG 1453  
Db 1585 ATAGTGTGGAAGTATCCCAATCAATGCAACAAATTTGCAAAATGACAGGTCATACATAGA 1644  
Qy 1454 GTTCTCTATCTTGGCATCTCTCCAGATGACAGACTATTGTAATCTGGAGCTGGAGATGAA 1513  
Db 1645 GTTTTATATTTGCTGTTGCTCTGATGTGAAACAACTCGTTACTTGTGCTGCTGATGAG 1704  
Qy 1514 ACGTTAGTCTTGGAAATGTTTT 1536  
Db 1705 ACCTTACGATTTTGAATGTATT 1727

RESULT 3  
US-10-369-493-26005  
; Sequence 26005, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 26005  
LENGTH: 1312  
TYPE: DNA  
ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-26005

Query Match 13.8%; Score 277; DB 12; Length 1312;  
Best Local Similarity 56.7%; Pred. No. 1.9e-59;  
Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

QY 636 TTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTGCAGATGATTTTATCTCA 695  
DB 372 TTCAAAACTCCATATAAATCTTGGATGCACCTTACTTAAAAATGACTTTTATTA 431  
QY 696 ATCTGGTAGATTGGTCTTACAAATGTTGGTGTGGTAACTGTGCTATC 755  
DB 432 ACTTGTGACTGGGCCAATCTAATGCTTGGCAGTCGACTAGCAAGCAGCATTTATC 491  
QY 756 TCTGGATGCTTGTAGCAGCAAGGTAACTAAATATATGATTTGGGGTTGATTTGTG 815  
DB 492 TCTGTGACGACGCAAGTGGTAAAGTGTTCATATACATGATTCGGAGCAACAAATCATG 551  
QY 816 TTTGTTCTTGTGGTGGCTCAACGTGTACTCTCTTGTCTTGGTAACTAACTAGTGA 875  
DB 552 TCACAAAGCGTCTTATGACAGGTAAAGNACTCAATAGCTGTAGGACAGCATTCAGAG 611  
QY 876 AAGTTTCAGATTTGGGATCAGCAAGATGCAAGAGATAAGATCAATGAGGGCCATCGGT 935  
DB 612 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 936 TACGTGTGGGGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 995  
DB 672 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 996 ATATTATCAACGAGATATAGCCACCAAGAGATTTTGTAGTAACTGTGAGACACA 1055  
DB 732 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 1056 AATCAGAGGTTTGTGGACTGAAGTGGTCAATATGATTAACCGTGGATCTGGAGGAA 1115  
DB 792 AACAGAAATATGCGATTGCAATGGATAGAGTCTAGCAGCTTGTCTTCAAGGAGGA 851  
QY 1116 ATGACAAATAATTTGTTTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTG 1175  
DB 852 ATGACAAATAATTTGTTTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTG 911  
QY 1176 AGCACAGCAGCTGTTAAAGCTATTTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1235  
DB 912 AGCATACAGACCGGTCAAGCAATTTGGATGGAGCCCTCAACAGAGGATCTCTGCA 971  
QY 1236 CTGGAGGAGGAACTGCAGATAGATGATTCGTTTGGATATCAACCAAACTCAACC 1295  
DB 972 CGCGCGGGGTACCATCGACAGATGCTTAACGATTCATAATCTTTGACTGGAAGGCTAC 1031  
QY 1296 TTACTGTATGACACTGGAGTCAAGTGTGCTATCTTGTCTGCTGCTGCTGCTGCTGCTGCT 1355  
DB 1032 AAAATAAGTTGATCTGGTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1091  
QY 1356 AACTAGTAAGCACATGGGTACTCCAGAACCGAGATTAATGTTTGGAGATACCCACCTA 1415

DB 1092 AAATGTTACTACTATGTTTGGAAAAAACCAAGTATCCTTATGGAATATCCTTCAT 1151  
QY 1416 TGTCAAAGCTGGGACTCTTTACCGGCCATCTTATAGGTTCTCTATCTTCCCATCTCTC 1475  
DB 1152 TGAATAATATCGCAATCTTACTTGCCTATCAAAACCGTGTCTCTTCTATGAGTC 1211  
QY 1476 CAGATGACAGACTTATGTAACCTGGAGCTGAGATGAACGCTTAGGTTCTTGGAACTTT 1535  
DB 1212 CCGATGCCAGTCTATAGTTACGGGAGCTGGGAGCAACGCTAAGATTTTGGAGCTTT 1271  
QY 1536 T 1536  
DB 1272 T 1272

RESULT 4  
US-10-369-493-26011  
Sequence 26011, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 26011  
LENGTH: 1312  
TYPE: DNA  
ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-26011

Query Match 13.8%; Score 277; DB 12; Length 1312;  
Best Local Similarity 56.7%; Pred. No. 1.9e-59;  
Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

QY 636 TTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTGCAGATGATTTTATCTCA 695  
DB 372 TTCAAAACTCCATATAAATCTTGGATGCACCTTACTTAAAAATGACTTTTATTA 431  
QY 696 ATCTGGTAGATTGGTCTTACAAATGTTGGTGTGGTAACTGTGCTATC 755  
DB 432 ACTTGTGACTGGGCCAATCTAATGCTTGGCAGTCGACTAGCAAGCAGCATTTATC 491  
QY 756 TCTGGATGCTTGTAGCAGCAAGGTAACTAAATATATGATTTGGGGTTGATTTGTG 815  
DB 492 TCTGTGACGACGCAAGTGGTAAAGTGTTCATATACATGATTCGGAGCAACAAATCATG 551  
QY 816 TTTGTTCTTGTGGTGGCTCAACGTGTACTCTCTTGTCTTGGTAACTAACTAGTGA 875  
DB 552 TCACAAAGCGTCTTATGACAGGTAAAGNACTCAATAGCTGTAGGACAGCATTCAGAG 611  
QY 876 AAGTTTCAGATTTGGGATCAGCAAGATGCAAGAGATAAGATCAATGAGGGCCATCGGT 935  
DB 612 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 936 TACGTGTGGGGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 995  
DB 672 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 996 ATATTATCAACGAGATATAGCCACCAAGAGATTTTGTAGTAACTGTGAGACACA 1055  
DB 732 TAAATTTACATCTGGGATATCGAGTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671  
QY 1056 AATCAGAGGTTTGTGGACTGAAGTGGTCAATATGATTAACCGTGGATCTGGAGGAA 1115



Db 792 AACAGAAATATCGGATTGCAATGGGATAGAACTTAGACAGCTTGTCTTCAGAGGGA 851  
Qy 1116 ATGACACAAATCTTTGTTGGAAATCAACACTCAACCCAGCCTGTCTCAAGTACTGTG 1175  
Db 852 ATGACAAATATTTATGCTATGGGATACCGCTCTTCAAGTCTTTACATTAAGTTGAAG 911  
Qy 1176 AGCAGACAGAGCTGTTAAAGCTATTGCAATGCTCTCTCATCTTCTTCAAGTCTTCTTGCAT 1235  
Db 912 AGCATACAGCAGCGGTCAAGCAATTTGGATGGAGCCCTCATCAACGAGGGATCTCTGCAA 971  
Qy 1236 CTGAGGAGGAACTGCAGATAGATGTTCTGTTTGGATACAAACACCAAACTCACACC 1295  
Db 972 GCGGCGGGGTACCATGACAGATGTTCAAGATCATATATCTTTGACTGGAGGCTAC 1031  
Qy 1296 TTAGCTGTATGACACTGGAAGTCAAGTTTGAATCTTTGCTGTCTGTCCTCAAAAATGTCAAAG 1355  
Db 1032 AAAATAAGTTGGATCTGCTCAAGTATGTAACATGCGCTGCTGCTCAAGACITCTTAAG 1091  
Qy 1356 AACTAGTAGACACACTGGAAGTCAAGTTTGAATCTTTGCTGTCTGTCCTCAAAAATGTCAAAG 1415  
Db 1092 AAAATGTTACTACTGCTTGTGCAAAACCAAGTATCCTTATGGAATATCTCTTCAT 1151  
Qy 1416 TGTCAAAGTGGGACTCTTACCGGCCATATCTATAGGTTCTCTATCTTGGCCATCTCTC 1475  
Db 1152 TGAATAATATGCCAATCTTACTGCCCCATACAAACCGGTGCTCTACCTTTCTATGAGTC 1211  
Qy 1476 CAGATGACAGACTATTGTTACTGAGCTGAGATGAACCGTTAGGTTCTGGAATGTTT 1535  
Db 1212 CCGATGGCCAGTCTATAGTTAGCGGAGCTGGGAGCGGAACCGTAAAGTTTGGAGCTTT 1271  
Qy 1536 T 1536  
Db 1272 T 1272

## RESULT 5

US-10-369-493-26013

; Sequence 26013, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 26013

; LENGTH: 1312

; TYPE: DNA

; ORGANISM: Schizosaccharomyces pombe

US-10-369-493-26013

Query Match 13.8%; Score 277; DB 12; Length 1312;

Best Local Similarity 56.7%; Pred. No. 1.9e-59;

Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

Qy 636 TTCCTCGATCGCCTTAAGGTTTGGATGACCTGCTTTGCAAGATGATTTTATCTGA 695

Db 372 TTCAAAAATCCATATAAATCTTGGATGACCTTACTTAAAAAATGACTTTTATTTAA 431

Qy 696 ATCTGGTATGTTGCTTTCACAAATGTTGGCTGTTGTTGGTAACTGTGCTATC 755

Db 432 ACTTGCTGCTGGGCGAATCTTAATGTTTGGAGTGGAGTACGACACATTTATC 491

Qy 756 TCTGGAATGCTTGTACGACGAGTAACATAATATGATTTGGGGTTGATTTG 815

Db 492 TCTGTGACGACAGCTGTAAGTGGTTCAATTACATGATTTCGGACCAACAATCATG 551  
Qy 816 TTTGTTCTGTTGGGCTCAACGCTGTACTCATCTTGTCTGTTGGAACTAAACAATGTA 875  
Db 552 TCACAAGCGTCTTATGGACAGGTAAGAACTCAATTAGCTGTAGGACAGATTCAGGAG 611  
Qy 876 AAGTTCAAGTTTGGGATGACGACAGATGCAAGAAGATGAATCAATGAGGCGCATCGT 935  
Db 612 TAATTTACATCTGGGATATCGAGTCTAGCAATCACTCAGATCTTTTAAAGTCTACTCG 671  
Qy 936 TAGCTGTCTGGGCGCTTGGCTGAGTTCATCTCTTTTGTCTTCTGTTGGACGGGATAAGA 995  
Db 672 AAGGGTGCAGCAITTAGCATGGAACGATAATACCTCACAAAGTGGAGCAAGACGAAG 731  
Qy 996 ATATTTATCAAGAGATATACCAACAGAGATTTTGTAGTAACTGTGAGACACA 1055  
Db 732 TAATTTTACATCATGACCTTAAGAGCACCGAGTTGCTGTGCCGAATGATGAAGTTCA 791  
Qy 1056 AATCAGAGGTTTGTGACCTGAAAGTGTCTATATGATAACCGTGAAGTTGSCATCTGAGAA 1115  
Db 792 AACAGAAATATGCGGATTTGCAATGGATAGAACTAGGACAGCTTGTCTTCAGGAGGA 851  
Qy 1116 ATGACAAAAATTTGTTTGGATCAACACTCAACCCAGCCTGCTCTCAAGTACTGTG 1175  
Db 852 ATGACAAATATTTATTCGTATGGGATTAACCGCTCTTCAAGTCTTTACATAAGTTTGA 911  
Qy 1176 AGCACACAGCAGCTGTAAAGCTATTGCAATGCTCTCTCATCTTCAAGTACTTCTTGCAT 1235  
Db 912 AGCATACAGCAGCGGTCAAGCAATTTGGATGAGCCCTCATCAACGAGGATCTCTGCAA 971  
Qy 1236 CTGAGGAGGAACTGCAGATAGATGTTCTGTTTGGATGACCAACCAAACTCACACC 1295  
Db 972 GCGGCGGGGTACCATGACAGATGTTCTAAACGATTCATAATATCTTTGACTGGAAGCTAC 1031  
Qy 1296 TTAGCTGTATGACACTGGAAGTCAAGTTTGAATCTTGTCTGTTCCAAAAATGTCAAG 1355  
Db 1032 AAAATAAGTTGGATCTGCTCAAGTATGTAACATGCGCTGCTTCAAGACTTCTTAAG 1091  
Qy 1356 AACTAGTAGACACAGATGGTACTCCAGAACCAAGATTAATGTTGGAGATACCCACTA 1415  
Db 1092 AAATTTACTACTCATGTTTGTCAAAAAACCAAGTATCTTATGGAATATCTCTTCAT 1151  
Qy 1416 TGTCAAAGTGGGACTCTTACCGGCCATATCTATAGGTTCTCTATCTTGGCATCTCTC 1475  
Db 1152 TGAATAATATGCGCAATCTTACTGCCCCATACAAACCGTGTCTCTTACTTCTAGTC 1211  
Qy 1476 CAGATGACAGACTATTGTTACTGAGCTGAGATGAACCGTTAGGTTCTGGAATGTTT 1535  
Db 1212 CCGATGGCCAGTCTATAGTTAGCGGAGCTGGGAGCGGAACCGTAAAGTTTGGAGCTTT 1271  
Qy 1536 T 1536  
Db 1272 T 1272

## RESULT 6

US-10-369-493-36526

; Sequence 36526, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 36526  
; LENGTH: 1116  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
; US-10-369-493-36526

Query Match 13.2%; Score 264.6; DB 12; Length 1116;  
Best Local Similarity 57.7%; Pred. No. 2.4e-56;  
Matches 542; Conservative 0; Mismatches 349; Indels 48; Gaps 2;

```
Qy 647 CTTTAAAGTTTGGAGCAGCTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGAT 706
    |||||
Db 178 CCCTACAAGTTCTCGATGCCAGACTTTCAGGACGATTTTACCTGAATTTGGTTGAT 237
    |||||
Qy 707 TGGTCTTACACAAATGTTGGCTGTTGGTTTGGGTAACCTGCTCTATCTCTGGAATGCT 766
    |||||
Db 238 TGGGGAGTAGTAGTTCTAGCGTTGGCTTAGGAACTCGGTGTACATGTGGAACTCA 297
    |||||
Qy 767 TGTAGCAGCAAGTAACAAATATATGATTTTGGGGTTG----- 806
    |||||
Db 298 CAAACTGGGAGGTTTACGAAATTTGTGAGCTTAAAGGATGACACAGCTCAGCGCTCACT 357
    |||||
Qy 807 -----ATGATTTGTTTGTCTTCTGTTGGTGGCTCAAC 839
    |||||
Db 358 TGGATCAAAAGGTAAGCGAGCTTTATGAAATGTTCTGGAGCTTATGTTGATAACATC 417
    |||||
Qy 840 -GTGCTACTCATCTGCTTGGAACTAACAAATGGTAAAGTTTCAGATTTGGGATGAGCA 898
    |||||
Db 418 TAGGGTACGCACCTTTCAAATGGTACGGGAAAGGTATGTTGCAAAATATGGGATGCAGAG 477
    |||||
Qy 899 AGATGCAAGAAGATAAGATCAATGAGGAGCCATCGGTAGCTGTCGGGCGCTTGGCCTGG 958
    |||||
Db 478 CGCTGTCGCGCCCTACGAGCATATGCGGACACCAATCGCTAGGGCGGTTGGCTTGG 537
    |||||
Qy 959 AGTTCACTCTTTTGTCTTCTGCTGGACGGGATAAGAAATTTATCAACAGATATACGC 1018
    |||||
Db 538 AACGATCATATCTGACATCGGCTCTCGGATCGGCATATTTTTCATCGTACGTCGG 597
    |||||
Qy 1019 ACACAAGAAGTTTGTAGTAAGTCTCAGGACACAAATCAGAGCTTTGTGACGTGAAG 1078
    |||||
Db 598 TCTCCTGACCAAGTATCTTCTGCACTTCTGCTCATAAGCAGGAAGTGTGGGCTCAGG 657
    |||||
Qy 1079 TGGTCATATGATAACCGTAGTTGGCATCTGGAGGAAATGACAAACAAATGTTGTTTGG 1138
    |||||
Db 658 TGGHAACGAGAGTGTCACTGGCATCAGGGGCAACGACAAATAGCTCTCTGTTTGG 717
    |||||
Qy 1139 AATCAACACTCAACCCAGCTGTCCTCAAGTATCTGTGAGCACACAGCAGCTGTTAAAGCT 1198
    |||||
Db 718 GACAAATTGAACGAGAGCCCTCTTTATCGCTTCTCAGACCATAACCGGCGCGTGAAGCT 777
    |||||
Qy 1199 ATTGCATGGTCTCTCATCTTTCATGGAATCTTCTGATCTGAGGAGAACTGACAGATAGA 1258
    |||||
Db 778 ATCAATGGTCAACCCATCAACACCACTTACTCGCATCCGGGGAGGATCTGCTGATCGA 837
    |||||
Qy 1259 TGTATTGTTTTTGGAAATACAAACCAAACTCACACCTTAGCTGTATGGACACTGGAAGT 1318
    |||||
Db 838 ACGATTAAATTTTGGACACGGCAGCGGTTCTCTTATCAAGGAGGTTGATACCGGAGC 897
    |||||
Qy 1319 CAGGTTTCATCTTGTCTGTCGCAAAATGTCACGAACTAGTAAAGCACACATGGGTAC 1378
    |||||
Db 898 CAAGTCTGTAACTTGGGATGTCGAAAGATTTCTGATGAATATCAGTACGATGGCTAT 957
    |||||
Qy 1379 TCCAGAACCAAGATTTATTTGTTGGAGATACCCCACTATGTCAAAGCTGGCGACTTTACC 1438
    |||||
Db 958 AGTCAGAACCAAAATTTGTCATCTGGAAGTATCTCTGATGGAGCAGATTTGTGCTTACG 1017
    |||||
Qy 1439 GGCCATCTTATAGGTTTCTATCTTGGCACTCTCCAGATGACAGACTATTTGTAAT 1498
    |||||
Db 1018 GGCCATCTTATCTGTTGCTCTATCTAGCCATGAGCCCGGACGCGGACGAGTGGTAACG 1077
    |||||
Qy 1499 GGAGCTGAGATGAACCGCTTAGGTTCTGGAATGTTTTC 1537
    |||||
Db 1078 GGTCCCGGCGAGACCCCTACGGTTCTGGAAGATATTC 1116
```

## RESULT 7

US-10-128-714-1038  
; Sequence 1038, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Broshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1038

; LENGTH: 1398

; TYPE: DNA

; ORGANISM: Aspergillus fumigatus

; US-10-128-714-1038

Query Match 13.2%; Score 263.8; DB 15; Length 1398;

Best Local Similarity 56.3%; Pred. No. 4.3e-56;

Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

```
Qy 609 ATAGTCCGGTTAAGGCTCCTAGGAAGTTCCCTCATCGCCTATATAGGTTTGGATGCAC 668
    |||||
Db 341 AGACTCTCCGCAAGCAGCCTCGCTAATAAGTACCCCTATAAGGTTCTCGACGCAC 400
    |||||
Qy 669 CTGCTTTGCAAGATGATTTTATCTGAACTCTGTAGATTTGCTTTCACAATGTTGG 728
    |||||
Db 401 CGGATTTGCAAGATGATTTTCTACTTGAATCTGTGCTGGGCGAGTAGTAATGTCTAG 460
    |||||
Qy 729 CTGTTGGTTTGGGTAACCTGCTCTATCTCTGGAATGCTTTGAGCAGCAAGTAATAAT 788
    |||||
Db 461 GTGTTGGGCTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGGACGGGTGACGAAAC 520
    |||||
Qy 789 TATGTGATTTGGGGGTTGATGATTG----- 813
    |||||
Db 521 TTTGCAACTAAGAGATGATCTGTCACAAGTTTAGCTGGATTCAGAGGGTAATGAGAT 580
    |||||
Qy 814 -----TGTTGTTCTGTTGGTGGGCTCAACGTGGTACTC 848
    |||||
Db 581 TTCTCTATATGTTGCAATGATTATTATGTTGTTTCACTGACTGACATTTGGACAGGGCACAC 640
    |||||
Qy 849 ATCTTCTGCTTGGAACTAACAATGCTAAAGTTTCAGATTTGGGATTCAGCAAGATCAAGA 908
    |||||
Db 641 ATCTTTTCAATAGAACTGGCAAGGTCTTTTACAGATATGGATTCGAGCATTTGTCGCC 700
    |||||
Qy 909 AGATAAGATCAATGGAGGGCCATCGGTTACGTTGCGGGCCCTTGGCTGGAGTTTCATCTC 968
    |||||
Db 701 GTCTTCGGACAATGATTGGGCATACCAATCGTGTGGGGCTCTTCTTGGAACTCATATA 760
    |||||
Qy 969 TTTTGTCTCTCGTGGACGGGATAGAATATTATCAACGAGATATAGCACAAGAG 1028
    |||||
Db 761 TCTTTACGTCAGGTTCTCGGGATCGACTGATCTTCCACCGTGATGTTCTGTTCCCGAGATC 820
    |||||
Qy 1029 ATTGTTTAGTAAATGTCAGGACACAAATCAGAGGTTTGTGGACTGAATGGTGCATATG 1088
```

Db 821 AGTATTACGTCGACTGTCGCGCCATAAGCAGGAAGTTTGGGACTCCGCTGGAACACCG 880  
Qy 1089 ATAAACCGTAGTGGCTCTCGAGGAATAGCAACAAATTTGTTGTTGGAATCAACACT 1148  
Db 881 AAGATGGCCAACTAGCTTCAGCGGAAATGACAAACAACTCATGGTGTGGGACAAAGCTCA 940  
Qy 1149 CAACCCAGCGTCTCCTCAAGTACTGTGAGCACACAGCAGCTGTTTAAAGCTATTGCAATGGT 1208  
Db 941 ATGAGACACCCGCTTATCGCTTTCGGACCATACCGCAGCGCTGAGCGCCATGCGCTGGT 1000  
Qy 1209 CTCCTCATCTTCATGGAATCTTTCATCTCGAGGAGGAATGTCAGATPAGATGATTGCTT 1268  
Db 1001 CACCCCATCAACACCATCTCTCGCTCAGGTGAGGACCGCGGACCGACCATCAAT 1060  
Qy 1269 TTTGGATACAAACCAACACTCACACCTTAGCTGTATGACACATGGAAGTCAGTTTGCA 1328  
Db 1061 TCTGGAACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACCGGTAGCCAGGTGTA 1120  
Qy 1329 ATCTTGCTGCTGCAAAATATGCAACGAATAGTAAGCACACATGCGTACTCCAGAAC 1388  
Db 1121 ACCTGGCATGTTCGAGAACTCGGACGAAATCATCAGTACACACGGCTACAGTCAGAAC 1180  
Qy 1389 AGATTATTGTTGGAGATACCCCACTATGTCAAAGCTGGGACTCTTACCGGCCATCTT 1448  
Db 1181 AGATTGTAATCTGGAATACCCGCGCATGGAGCAGATTGTATCGTCAACCGGCCACACTT 1240  
Qy 1449 ATAGGTTCTCTATCTTCCCATCTCTCCAGATGACACAGACTATTGTAACCTGGAGCTGGAG 1508  
Db 1241 TCCGTTCTCTACTCTTGGATGAGCCCGCATGGAGCAGACCGTGTATCCCGAGCTGGT 1300  
Qy 1509 ATGAAACGCTTAGTTCTGGAATGTTTTC 1537  
Db 1301 ACGAAACGTTGAGATTCTGGAAGATTTTC 1329

RESULT 8.  
US-10-128-714-2038  
; Sequence 2038, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Weng  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Broshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; PRIOR FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2038  
; LENGTH: 1398  
; TYPE: DNA  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-2038

Query Match 13.2%; Score 263.8; DB 15; Length 1398;  
Best Local Similarity 56.3%; Pred. No. 4,3e-56;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

Qy 609 ATATGCCGTTAAGGCTCCTAGGAGGTTCTCGATCGCCCTTATAAGGTTTTGGATGCAC 668  
Db 341 AGACTCTCCGACAGCAGCTCTGCTACGTTAATAAGTACCTCTAAGGTTCTCGACGAC 400  
Qy 669 CTGCTTTGCAAGATGATTTTTATCTGAATCTGGTAGATTGGTCTTCAACAAATGTTGG 728  
Db 401 CCGATTTGCAAGATGATTTTACTTGAATCTGGTAGATTGGTCTTCAACAAATGTTGGTAG 460  
Qy 729 CTGTTGGTTTGGGTAACCTGCTCTATCTCTGAAATGCTTTAGCAGCAAGGTAACATAAT 788  
Db 461 GTGTTGGGCTAGGCAATTCAGTGTATCATGTGAAATTCGAAATCCGAGCGGTGACGAAC 520  
Qy 789 TATGTGATTTGGGGCTTGATGATTG-----TGTGTTCTGTTGGTGGCTCAACGTTGATCTC 813  
Db 521 TTTTGGAACTAAGACATGATCTGTCAACAAGTGTAGCTGGATTTCAGAGGGTAATGACAT 580  
Qy 814 -----TGTGTTCTGTTGGTGGCTCAACGTTGATCTC 848  
Db 581 TTTCTCTATATGTTGGCAATGATTAATATGTTTTCATGACTGACATTTGGACAGGGCACAC 640  
Qy 849 ATCTTGGCTGTTGGAATAACAAATGTAAGTTTCAATTTGGGATGACAGCAAGATGCAAGA 908  
Db 641 ATCTTTCAATAGAACTGGAAGGCTCTGTACAGATATGGATGGATGAGAGCATTTCTGCC 700  
Qy 909 AGATAAGATCAATGAGGGCCATTCGGTTACGTTGTCGGGGCTTGGCTTGGAGTTTCTCTC 968  
Db 701 GTCTTGGCAATGATTTGGGCATCAACATCGTGTGGGGCTCTTCTTGGAAACGATCATATA 760  
Qy 969 TTTTGTCTTCTGCTGGACGGGATAAGAATATTTATCAACGAGATATACGCAACAAGAG 1028  
Db 761 TCTTACGTCAGTTCTCGGATCGACTGATCTCCACCGTGATGTCTGTTCCCGCAGATC 820  
Qy 1029 ATTTTGTAGTAACTGTGACAGACAAATTCAGAGGTTTGTGGACTGAGTGGTCTATATG 1088  
Db 821 AGTATTTACGTCGACTGTCTCGGCCATAAGCAGGAAGTTTGGGACTCCGGTGGAAACACG 880  
Qy 1089 ATAAACGCTGAGTTGGCATCTGGAGGAATGACAAATTTGTTTGGATCAACACT 1148  
Db 881 AGATGGCCCACTAGCTTCAGCGGGAATGACAACTCATGCTGTTGGGACAGCTGA 940  
Qy 1149 CAACCCAGCGCTGTCTCAAGTACTGTGAGCAACAGCAGCTGTTTAAAGCTATTGCAATGGT 1208  
Db 941 ATGAGACACCGCTTTTATCGCTTCTCGGACCATACCGCAGCCGTGAAGGCCATCGCTGGT 1000  
Qy 1209 CTCTCATCTTCATGAGCTTCTTTCATCTGGAGGAGGAACCTGCAGATAGATGATTGTT 1268  
Db 1001 CACCCCATCAACACCATCTCTCGCTCAGGTGGAGGACAGCGGACCGGACCATCAAT 1060  
Qy 1269 TTTGGAATACAAACCAAACTCACACCTTACCTGTATGAGCACTGGAAGTCAGTTTGA 1328  
Db 1061 TCTGGAACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTA 1120  
Qy 1329 ATCTTGTCTGGTCCAAATGTCAGGAAGTATGAGCAACATGAGTACTCCAGAAC 1388  
Db 1121 ACCTGGCATGGTCAAGAACTCGGAGGAATCATCTGTAACAACGGCTACAGTCAAGAAC 1180  
Qy 1389 AGATTATTGTTGGAGATACCCCATCTATGTCAAAGCTGGCGACTCTTACCGGCCATCTT 1448  
Db 1181 AGATTGTAATCTGGAATACCCGCGCATGGAGCAGATTGTATCGCTGACCGGCCACACTT 1240  
Qy 1449 ATAGGTTCTCTATCTTGGCACTCTCCAGATGACAGACTATTGTAACCTGGAGCTGGAG 1508  
Db 1241 TCCGTTCTCTACTCTTGGATGAGCCCGGATGGAGACCGTGTATCCCGAGCTGGT 1300  
Qy 1509 ATGAAACGCTTAGGTTCTGGAATGTTTTC 1537  
Db 1301 ACGAAACGTTGAGATTCTGGAAGATTTTC 1329

RESULT 9  
US-10-128-714-7038  
; Sequence 7038, Application US/10128714

Publication No. US20030119013A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Bo  
APPLICANT: Hu, Wengqi  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Eroshekin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
TITLE OF INVENTION: Methods of Use  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
CURRENT FILING DATE: 2002-04-23  
PRIOR APPLICATION NUMBER: US 60/285,697  
PRIOR FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: US 60/287,066  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/295,890  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/303,899  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/316,362  
PRIOR FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7038  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-7038

Query Match 13.2%; Score 263.8; DB 15; Length 1875;  
Best Local Similarity 56.3%; Pred. No. 5.1e-56;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;  
QY 609 ATAGTCGGTTAGGCTCTAGGAAGGTCCTCGATCGCTTATAGAGTTTGGATGCAC 668  
DB 818 AGACTCTCGAAGAGCCTCGTACGTTAATAAGTACCCATATAGGTTCTCGACGCAC 877  
QY 669 CTGCTTGGCAAGATGATTTTATCTGAATCTGTAGATGTTCTTACCAATGTTGG 728  
DB 878 CCATTTGCAAGATGATTTTATCTGAATCTGTAGATGTTCTTACCAATGTTGG 937  
QY 729 CTGCTTGGTAACTGTGTCTATCTCTGGAATGCTTGTAGCAGCAAGGTAATCAAT 788  
DB 938 GTGTGGCTTAGCAATTCAGTGTACATGTGGAATTCGAATACCGACGGGTGACGAAC 997  
QY 789 TATGTGATTTGGGGTGTGATTTG-----TGTTTGTCTGTTGGTGGGCTCAACGTGTA 813  
DB 998 TTTGCGAACTAAGATGATGATCTGTACAAAGTGTAGCTGATTCAGAGGTAATGAGAT 1057  
QY 814 -----TGTTTGTCTGTTGGTGGGCTCAACGTGTA 848  
DB 1058 TTCTCTATATGTTGGCAATGATTTATGTTGTTTCAATGACTGACATTGGACAGGCACAC 1117  
QY 849 ATCTTGTGTTGAACTAAAGTAAAGTTTCAAGTTTGGATGGGATGAGATGCAAGA 908  
DB 1118 ATCTTTCAATAGGAATGGAAGGTTCTGTACAGATATGGGATGAGATGTCGCGC 1177  
QY 909 AGATAAGATCAATGGAGGCGCATCGTTACGTGTGCGGGCTTTGGCTGGAGTTCAATCTC 968  
DB 1178 GTCTTCGCAATGATTTGGGCATACCAATCGTGTGGGGCTCTTGTGGAACGATCAT 1237  
QY 969 TTTTGTCTTCTGGTGGAGCGGATAGAAATTTATCAACAGATATACGACCAAGAG 1028  
DB 1238 TCCTTAGCTCAGGTTCTCGGGATCGACTGATCTTCCACCGTGTATGTTCTTCCCGAGATC 1297  
QY 1029 ATTTTGTAGTAACTCTCAGGACACAAATCAGAGGTTTGTGACTGAAATGTCATATG 1088  
DB 1298 AGTATTACGTGCACTGTCCGGCCATAGCAGGAAGTTTGGGACTCCGGTGAACCG 1357  
QY 1089 ATACCGTGTAGTGGCACTGTGGAGGAATGACAAATGTTTGTGTAATCAATCT 1148

DB 1358 AAGATGGCCAACTAGCTTTCAGGGCGGAATGACAACTCATGTTGGGCAAGCTGA 1417  
QY 1149 CAACCCAGCTGTCTCTCAAGTACTGTGAGCACACAGCAGCTGTAAAGCTATTTCATGGT 1208  
DB 1418 ATGAGACACCGCTTTATTCGCTTCTCGGACCATACCGCAGCCGTGAAGGCCATCGCTGGT 1477  
QY 1209 CTCTCATCTTCATGGACTTCTTGCATCTCGAGGAGAACTGCAGATAGATGATGTTGTT 1268  
DB 1478 CACCCCATCAACCATCTCTCTCGCTCAGGTGGAGGACCGGGGACCGGACCAATCAAT 1537  
QY 1269 TTTGGAATACAAACCAACAACTCACACCTTAGCTGTATGAGACACTTGAAGTCAGGTTTSCA 1328  
DB 1538 TCTGGAACACTGCTACTGTTCTACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1597  
QY 1329 ATCTTGTCTGTCCTCAAAAATGTCAACGAATAGTAAAGACACATAGCTGTTCCAGAAC 1388  
DB 1598 ACCTGGCATGTCGAGAACTCGGACGAATATCATAGTACACAGGGTACAGTCAGAAC 1657  
QY 1389 AGATTATTGTTGGAGATACCCACTATGTCAAAGCTGGGCACTCTTACCGGGCCATACTT 1448  
DB 1658 AGATTGTAATCTGGAATACCGGCGCATGGAGAGATTGTATCGCTGACCGGCCACACTT 1717  
QY 1449 ATAGGGTCTCTATCTTGGCATCTCTCCAGATGAGACAGACTATTGTAACTGGAGCTGGAG 1508  
DB 1718 TCCGTGTTCTCTACCTTGGCATGAGCCCGATGACAGCCGTCGTTTACCGGAGCTGGT 1777  
QY 1509 ATGAACGCTTAGGTTCTGGAATGTTTTC 1537  
DB 1778 ACGAAGCTTGAGATTCTGGAAGATTTC 1806

RESULT 10  
US-10-128-714-6038  
Sequence 6038, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Bo  
APPLICANT: Hu, Wengqi  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Eroshekin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
TITLE OF INVENTION: Methods of Use  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
CURRENT FILING DATE: 2002-04-23  
PRIOR APPLICATION NUMBER: US 60/285,697  
PRIOR FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: US 60/287,066  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/295,890  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/303,899  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/316,362  
PRIOR FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6038  
LENGTH: 1993  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-6038

Query Match 13.2%; Score 263.8; DB 15; Length 1993;  
Best Local Similarity 56.3%; Pred. No. 5.3e-56;  
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;  
QY 609 ATAGTCGGTTAAGGCTCTAGGAAGGTTCTCGATCGCTTATAGAGTTTGGATGCAC 668  
DB 936 AGACTCTCGCAAGAGGCTGCTACGTTAATAAGTACCCATATAGGTTCTCGACGCAC 995

669 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 728  
Db CCGATTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 1055  
729 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 768  
Db GTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 1115  
789 TATGTGATTTGGGGTGTGATGTTG----- 813  
1116 TTTGCGAACTAAGAGATGATGATGTCACAAAGTGTAGCTGATTTAGAGGGTATGAGAT 1175  
814 -----TGTTTCTCTGTTGGTGGCTCAACGGTACTC 848  
1176 TTCTCTATATGTTGGCAAGATGATTTATGTTTCTGATGATGATGATGATGATGATGAT 1235  
849 ATCTTGTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 908  
1236 ATCTTGTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295  
909 AGATTAAGATCAATGAGAGGCAATCGTTACGTTGCGGGCTTGGCTGGAGTTCACTC 968  
1296 GTCTTTCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355  
969 TTTTGTCTTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028  
1356 TCTTACGTCAGTTCTCGGATCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415  
1029 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088  
1416 AGTATTTACGTCAGTTCTCGGATCGATGATGATGATGATGATGATGATGATGATGATGAT 1475  
1089 ATAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
1476 AGATGCGCACTAGTTCTCGGATCGATGATGATGATGATGATGATGATGATGATGATGAT 1535  
1149 CAACCCGCTGCTCCTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGATGATGATGAT 1208  
1536 ATGAGACACCGCTTTATCGCTTCTCGGATCGATGATGATGATGATGATGATGATGATGAT 1595  
1209 CTCTCTCATCTTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGATGATGATGATGAT 1268  
1596 CAACCCGCTGCTCCTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGATGATGATGAT 1555  
1269 TTTGGAATCAACCAAACTCCTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGATGAT 1328  
1656 TCTGGAACACTGCTACTGCTTCTCGGATCGATGATGATGATGATGATGATGATGATGATGAT 1715  
1329 ATCTTGTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388  
1716 ACCTGCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775  
1389 AGATTAATGTTGGAGATACCCCACTATGTCACAAAGTGTGATGATGATGATGATGATGAT 1448  
1776 AGATTTGTAATCTGGAATACCGCGCATGGAGGATGATGATGATGATGATGATGATGATGAT 1835  
1449 ATAGGGTCTCTATCTTGGCACTCCTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGAT 1508  
1836 TCCGTGTTCTTACCTTGGATGAGCCCGATGAGCAACAGCAAGTGTGATGATGATGATGATGAT 1895  
1509 ATGAAACGCTTAGGTTCTGGAATGTTTC 1537  
1896 ACGAAACGTTGAGATCTGGAAGATTTTC 1924

## RESULT 11

US-10-128-714-38  
; Sequence 38, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wendi  
; APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos  
APPLICANT: Broshkin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
TITLE OF INVENTION: Methods of Use  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
CURRENT FILING DATE: 2002-04-23  
PRIOR APPLICATION NUMBER: US 60/285,697  
PRIOR FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: US 60/287,066  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/295,890  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/303,899  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/316,362  
PRIOR FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 38  
LENGTH: 3398  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-38

Query Match 13.2%; Score 263.8; DB 15; Length 3398;

Best Local Similarity 56.3%; Pred No. 7.2e-56;

Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGGTTAAGGCTCTAGAGAGGTTCTCGATCGGCTTATAGGTTTGGATGAC 668

Db 1341 AGACTCCTCGCAAGAGCGCTCGCTACGTTAATAAAGTACCCCTATAAGGTTCTCGACGCAC 1400

QY 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 728

Db 1401 CCGATTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 1460

QY 729 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGCTCTTCAACAATGTTGG 788

Db 1461 GTGTTGGCTTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGAGCGGTTGACGAAAC 1520

QY 789 TATGTGATTTGGGGTGTGATGTTG----- 813

Db 1521 TTTGCGAACTAAGAGATGATGATGTCACAAAGTGTAGCTGATTCAGAGGGTATGAGAT 1580

QY 814 -----TGTTTGTCTGTTGGTGGCTCAACGTTGATCTC 848

Db 1581 TTCTCTATATGTTGGCAATGATTTATGTTGTTTATGATGATGATGATGATGATGATGATGAT 1640

QY 849 ATCTTGTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 908

Db 1641 ATCTTGTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1700

QY 909 AGATAAGATCAATGAGAGGCTCGGTTACGTTGCGGGGCTTGGCTGGAGTTTCACTCTC 968

Db 1701 GTCTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1760

QY 969 TTTTGTCTTCTGTTGGCACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028

Db 1761 TCCTTACGTCAGGTTCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1820

QY 1029 ATTTTGTAGTAACTGTCAGGACCAAAATCAGAGGTTTGTGCACTGGAAGTGTCTCATATG 1088

Db 1821 AGTATTTACGTCAGTGTCTCGGCTTACGAGAGGTTTGGGCTCGGCTCGGCTCGGCTCGG 1880

QY 1089 ATAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148

Db 1881 AAGATGCGCAACTAGCTTTCAGGCGGAAATGACAACTCATGTTGTTGGGCAAGCTGGA 1940

QY 1149 CAACCCGCTGCTCTCAAGTACTGAGCAACAGCAAGTGTGATGATGATGATGATGATGATGAT 1208

Db 1941 ATGAGACACCGCTTTATCGCTTCTCGGACCATACCGAGCGGTTGAGGCTCGCTGCT 2000

QY 1209 CTCCTCATCTTCATGGACTCTTTCATCTCGAGGAGAACTGCAGATAGATGATTTCGTT 1268  
DB 2001 CACCCATCAACACCATCTCTCGCTCAGTGGAGGACGGCGGACCCATCAAT 2060  
QY 1269 TTTGGAATCAACACAACTCAGCTTACCTGATGATGAGACATGGAAGTCAGTTGCA 1328  
DB 2061 TCTGGAACACTGCTACTCTGTTCTCATCTGATCAAGGAAGTTGATACGGGTAGCCAGTGTGA 2120  
QY 1329 ATCTTGCTGCTCCAAATGTCAACGAAGTCTAGTAAGCACAACATGGGTACTCCCAAGAC 1388  
DB 2121 ACCTGGCATGTCGAGAACTCGGAGGAATCATCAGTACACAGGCTACAGTCAGAAC 2180  
QY 1389 AGATTATTTGTTGAGATACCCCACTATGTCAAAGCTGGGCACTCTTACCGGCCATCTT 1448  
DB 2181 AGATTGTAATCTGGAATACCCGCAATGGAGCAGATTTGATCGCTGACCGGCCACACTT 2240  
QY 1449 ATAGGGTTCTATCTTGGCATCTCTCCAGATGACAGACTATTTGTAACCTGGAGCTGGAG 1508  
DB 2241 TCCGTGTTCTTACCTTGGATGAGCCCGATGAGACAGCGTCTGTACCGGAGCTGGTG 2300  
QY 1509 ATGAACGCTTAGGTTCTGGAATGTTTTTTC 1537  
DB 2301 ACGAAGCTTGAGATTCTGGAAGATTTTC 2329

## RESULT 12

US-10-128-714-5038

; Sequence 5038, Application US/10128714

; Publication No. US20030119013A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Eroshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and

; FILE OF INVENTION: Methods of Use

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5038

; LENGTH: 3993

; TYPE: DNA

; ORGANISM: *Aspergillus fumigatus*

US-10-128-714-5038

Query Match 13.2%; Score 263.8; DB 15; Length 3993;

Best Local Similarity 56.3%; Pred. No. 7.9e-56;

Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCCGGTTAAGGCTCTAGGAGGTTCTCGATCCCTTATAAGTTTGGATGCAC 668

DB 1936 AGACTCCTCGCAAGCAGCCTCGTACGTTAATTAAGTACCCTATTAAGTTCTCGACGCAC 1995

QY 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGATGTTGTTCTTACACAACTGTTGG 728

DB 1996 CCGATTGCAAGATGATTTTACTTACTTCTGATCTGGTGGACTGGGCGAGTAGTAACTGCTAG 2055

QY 729 CTGTTGGTTGGTGAATGTTCTATCTCTGGAATGTTGTAGCAGCAAGGTAATAAAT 788

DB 2056 GTGTTGGCTTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGGAGCGGTGCGAAAC 2115  
QY 789 TATGTGATTTGGGGTGTGATGATTG----- 813  
DB 2116 TTTGGAATCAAGAGATGATCTGTCAAGTGTAGTTCAGAGGTAATGAGAT 2175  
QY 814 -----TGTGTTCTCTGTTGGTGGCTCAAGTGGTACTC 848  
DB 2176 TTCTCTATATTTGGCAATGATTTATGTTTTCATGACTGACATTTGGACAGGGCACAC 2235  
QY 849 ATCTTGCTGTTGGAACTAAACATGTTAAAGTTTCAAGTTTGGGATGCGAGAGATGCAAGA 908  
DB 2236 ATCTTTCAATAGAACTGCGAAAGTCTTGTACAGATATGGGATGCGAGCAATTTGCGCC 2295  
QY 909 AGATAAGATCAATGAGAGGCGCATCGGTTACCTGTGCGGGCCCTTGGCCCTGGAGTTCACTC 968  
DB 2296 GTCTTCGGACAAATGATTTGGGCATACCAATCGTGTGGGGGCTCTTCTTTGGAAACGATCATA 2355  
QY 969 TTTTGTCTTCTGTTGCGAGCGGATAAGATATTTTATCAACGAGATATATCGCACACAAGAAG 1028  
DB 2356 TCTTTACGTCAAGTTCTCGGATGAGTGTCTCCACCGTGTGTTGTTCCCGCAGATC 2415  
QY 1029 ATTTTGTAGTAACTGTGAGACACAATCAGAGGTTTGTGGACTGAAGTGGTCAATG 1088  
DB 2416 AGTATTTAGCTGACTGTCCGCCATAGCAGGAAGTTTGGGACTCCGGTGGAAACCG 2475  
QY 1089 ATAAACGTTGAGTTGGCATCTGGAGGAATGACAAATTTGTTGTTTGGAAATCAACT 1148  
DB 2476 AAGATGGCCACTAGCTTCAGCGGAATGACAAACTCATGTTGTGGACACAGCTGA 2535  
QY 1149 CAAACCCAGCTGTCTCTCAAGTACTGTGAGCACACAGCAGCTGTTTAAAGCTATTGCAATGGT 1208  
DB 2536 ATGAGACACCGCTTTTATCGTCTCTCGGACCATATACCGCAGCGGTGAAGGCCATCGCTGGT 2595  
QY 1209 CTCTCACTTCTCATGACTTCTTGCATCTGGAGGAGAACTGCGAGATAGATGATTCTGTT 1268  
DB 2596 CACCCCATCAACACCATCTCTCGCTCAGGTGGAGGACCGCGGACCGGACCATCAAT 2655  
QY 1269 TTTGGAATACAAACACAAACTCACACCTTAGCTGTATGACACACTGGAAGTCAAGTTTGA 1328  
DB 2656 TCTGGAACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 2715  
QY 1329 ATCTTGTTCTGGTCCAAAATGTCAACGAAGTATGTAAGCACACATCGGTACTCCGAGAAC 1388  
DB 2716 ACCTGGCATGTCGAAGAACTCGGACGAAATCATCAGTACACACCGCTACAGTCAAGAAC 2775  
QY 1389 AGATTATTTGTTGGAGATACCCCACTATGTCAAAAGCTGCGACTCTTACCGGCCATCTT 1448  
DB 2776 AGATTGTAATCTGGAATATACCCGCGCATGGACAGATTGTATCGCTGACCGCCACACTT 2835  
QY 1449 ATAGGTTCTCTATCTTGCCATCTCTCCAGATGAGCAGACTATTGTAAGTGGAGCTGGAG 1508  
DB 2836 TCCGTTTCTCTACTTTCGATGAGCCCGGATGAGCCCGATGAGACAGACCGTCTTACCGGAGCTGGTG 2895  
QY 1509 ATGAACGCTTAGGTTCTGGAATGTTTTTC 1537  
DB 2896 ACGAAGCTTGAGATTCTGGAAGATTTTC 2924

## RESULT 13

US-10-369-493-45642

; Sequence 45642, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B



; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 45642  
; LENGTH: 1701  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-45642

Query Match 11.2%; Score 225.6; DB 12; Length 1701;

Best Local Similarity 53.3%; Pred. No. 2.3e-46; Indels 6; Gaps 1;  
Matches 502; Conservative 0; Mismatches 434; Indels 6; Gaps 1;  
614 CCGGTTAAGGCTCTAGGAAGGTTCTCGATCGCCTTATAAGGTTTGGAGTCACTGCT 673  
718 CCGGGAAGCAATTAGACAAATAGCTAAAGTTCCTATAGATTTTGGATGCCCATCA 777  
674 TTGCAAGATGATTTTATCTCAATCTGGTAGATTTGCTTACACACATGTTGGCTGTT 733  
778 TTAGCAGACGATTTTATCTACAGCTTGATAGTTGGTCAAGTACTGAATTTTGGCAGTT 837  
734 GGTITGGGTAAGTCTGCTATCTATCTGGAATCTTGTAGCAGCAAGGTAACATAATATGT 793  
838 GCTCTTGGAAATCGATTTTAAACCGATAATAACACTGCGGACGTTGTTCACTTATGC 897  
794 GATTTGGGGTGTGATGATGTTGTTGTTGTTGGTGGCTCAACGTTGTTCTCATCTTT 853  
898 GACACGGAAACGAATAACAAGCT-----TAAGCTGGATTGGAGCAGGCTCTCATCTG 951  
854 GCTGTTGGAACATAAATGTTAAAGTTCAGATTTGGGATGAGCAAGATGCAAGAATA 913  
952 GCAGTAGTCAAGCAATGGACTTGTAGAAATTTATGACGTAATGAAAGAAATGATTT 1011  
914 AGATCAATGAGGGGCATCGGTTAGCTGCGGGCTTGGCTGGAGTTCACTCTTTTGG 973  
1012 AGGAGTTTGTGCGGACATATTGACAGAGTAGCGTGTATTCTTGGAAATATCATGTTTG 1071  
974 TCTTCTGTGGACGGGTAAGAATATTTATCAACAGATAACGCAACACAGAAAGATTTT 1033  
1072 AGCTTGGAGTAGAGATCTGCAATCTGATAGGATTTAGGATGCCAGATCCCTTT 1131  
1034 GTTAGTAACCTGTCAAGSACAAATCAGAGTTTGTGCACTGAAGTGTCTATGATAAC 1093  
1132 TTTGAAACTATAGAACTCGCACACTCAGGAAGTCTGTGGCTTAAAGTGAATGTAGCGGAC 1191  
1094 CGTAGTTGGCATCTGAGGAAATGACACAAATTTGTTTGGATCAACACTCAACC 1153  
1192 AACAACTTGCCTCAGGTGGTAAACGATACGTCGTCATGTTTACGAAGAACGTCGAA 1251  
1154 CAGCCTGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCAATGTCCT 1213  
1252 TCTCCAACTCTGACGTTTGACGAACTAAGGCTGCTGTTAAGGCAATGGCCTGCTCCT 1311  
1214 CATCTTCAGGACTTCTGATCTGAGGAGAACTGAGATAGTATGTTTCTGTTTGG 1273  
1312 CATAACGAGGTGTGCTAGTCTACCGGTGGTGATCAGCAGATAGAACTAAATCTGG 1371  
1274 AATACAAACCAAACTCAGCCTTATGATGAGCACTGGAAGTCAGGTTTGCATCTT 1333  
1372 AACGTTAATACGTCATTAAGATGATGATATAGATGCGGCTCTCAAAATATGATATG 1431  
1334 GTCTGGTCCAAAATGTCAGCAATAGTAAAGCACATGGTACTCCAGAACCAAGATT 1393  
1432 GTATGTCAGGAATACTAACGAGCTTGTGACATCACATGGTTACTCAAAATATAACTTA 1491  
1394 ATTGTTGGAGATACCCCACTATGTCAAAGCTGGGACTCTTACCGGCCATCTTATAGG 1453  
1492 ACCTTGGGAGCTGTAACTCTATGGATCCAAATTGCAATTCGAAAGGTCATAGTTTCAGG 1551  
1454 GTTCTCTATCTTGGCCATCTCTCCAGATGGACAGACTATTTGTAATCTGGAGCTGGAGTAA 1513

Db 1552 GTTCTGCATTTAAAGTTATCAATGATGAACCAAGGTTGTTTCTGGGCTGGAGATGAA 1611  
QY 1514 ACGCTTAGGTTCTGGAATGTTTCCCTTCCCTAAATCACAG 1555  
Db 1612 ACATTGGCATATTGGAACCTTTTGTATAAACCAAAAGCGAAG 1653

## RESULT 14

US-09-938-842A-988  
; Sequence 988, Application US/09938842A  
; Patent No. US20020160378A1

; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 988

; LENGTH: 1344  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-988

Query Match 10.6%; Score 213.6; DB 10; Length 1344;  
Best Local Similarity 54.2%; Pred. No. 2.2e-43;  
Matches 513; Conservative 0; Mismatches 409; Indels 24; Gaps 3;

QY 615 CGGTTAAGGCTCTAGGAAGGTTCTCGATCCCTTATAAGGTTTGGATGACCTGCTT 674  
Db 320 CTGTTAAGCTCTCGTGCATACATTCCTCAGACTTCTGAGAAACCTTGGATGACCTGACA 379  
QY 675 TGCAACATGATTTTATCTGAATCTGTAATGTTCTTCAACAATGTTGGCTGTTG 734  
Db 380 TTGTTGACGATTTCTACCTCACTTCTGCTGACTGGGAAGTCAATGCTTTAGCCATAG 439  
QY 735 GTTTGGGTAACTGTGCTATCTCTGGAATGCTTTAGCAGCAAGGTAATAATTATGTG 794  
Db 440 CGTTGGACCACACTGTCTACTTGTGGGATGCTTCCACTGTTCTACATCTGAGCTGTGA 499  
QY 795 ATTTGGGGTGTGATGTTG---TGTTTGTCTGTTGGTGGCTCAACGCTGCTACTCATC 851  
Db 500 CCATTGATGAGGAAGGACCTGTCACAAGTATCACTGGGCTCTGATGTCGTCATG 559  
QY 852 TTGCTGTTGGAACTAAACAATGGTTAAAGTTTCAAGTTTGGGA---TGCAAGCAAGATGCAAGA 908  
Db 560 TTGCACTTGGACTCAACAACCTCTGAAGTCCAGCTGTGGGATTTCTGCATCCAAACCGTCAAC 619  
QY 909 AGATAAGATCAATGGAGGCCCATCGGTTACGTTGCGGGCTTGGGCTCGAGTTCATCTC 968  
Db 620 TGAGAACATTTAAAGGGTGTCCACAGTCAAGAGTAGATCATCTGGCATGGAACATCAT 679  
QY 969 TTTTGTCTTCTGGTGACGCGGATAGAATATTTATCAACAGATATACGCAACAGAAG 1028  
Db 680 TCCTTACTACTGAGGAATGATGGACTGATCATCAACATGATGTGAGGATCAGATCAC 739  
QY 1029 ATTTGTTAGTAACTGTGAGGACCAATCAGAGGTTTGTGAGCTGAGTGTGTCATATG 1088  
Db 740 CCATTGTGGAACCTTACAGAGGTCACTCAAGAGGTTTGTGGGCTCAAGTGTCTAGGAT 799  
QY 1089 ATAACTGATTTGGCATCTGAGGAAATGACAAACAATTTGTTTGTGGAATC-----1142  
Db 800 CTGACAAACACTAGCAAGTGGTGGCAACGACAAATGTGTTGATACATCTGGATCGTTCTG 859

QY 1143 -----AACACTCAACCCAGCCTGCTCCTCAAGTACTGTGTGACACACAGCAGCTG 1190  
Db 860 TCGCTTCTCAAACTCAACACACAAATGGCTGCACAGGCTTGAGGACACATACATCTGCTG 919  
QY 1191 TTAAGCTATTGATGGTCTCTCATCTTCATGGAATCTCTTGCATCTGAGGAGGAACTG 1250  
Db 920 TGAAGGCTCTTGGTGGTGGCTTTCACAGCGAAATTTGCTTGCACCTGGTGGTGGAG 979  
QY 1251 CAGATAGATGATTTGTTTTTGGAAATACACACAACTCAACCTTTAGCTGTATGGACA 1310  
Db 980 GAGACAGGACGATAAGTTCTGGAATCTACACTGGGCTTCTTGAATTCAGTAGACA 1039  
QY 1311 CTGGAAGTCAAGTTTGCATCTTGTCTGTGTCCTCAAAATGTCACGAACTAGTAAGCACAC 1370  
Db 1040 CTGGTTCCCAAGTTTGTCTGTGTTATGGAGCAAGAAATGAAAGAGAGTTGCTTAGCTCAC 1099  
QY 1371 ATGGGTACTCCAGAACCAATATTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGA 1430  
Db 1100 ATGGGTTTACAGAAATCAGCTTACACTTTGGAGTACCCATCCATGGTGAATGGCTG 1159  
QY 1431 CTCTTACCGCCATCTATAGGTTCTCTATCTTCCCATCTCTCCAGATGGACAGACTA 1490  
Db 1160 AGCTCACTGGTCTACATCAAGAGTTCTATATATGCGCCAGAGTCCAGATGGTTGACTG 1219  
QY 1491 TTGTAAGTCAAGCTGGAGTGAACGCTTAGGTTCTGGATGTTTT 1536  
Db 1220 TAGCTTACGAGCAGGAGACGAGACCTTAGGTTTGGAACTGTTTT 1265

## RESULT 15

US-09-938-842A-988  
; Sequence 988, Application US/09938842A  
; Publication No. US2004009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 988  
; LENGTH: 1344  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-988

Query Match 10.6%; Score 213.6; DB 12; Length 1344;  
Best Local Similarity 54.2%; Pred. No. 2.2e-43;  
Matches 513; Conservative 0; Mismatches 409; Indels 24; Gaps 3;  
QY 615 CGGTTAAGGCTCTAGGAAGGTTCTCGATCGGCTTATAAGGTTTGGATGACACCTGCTT 674  
Db 320 CTGTAAAGCCCTCGTCGATACATTCCTCAGACTTCTGAGAGAACCTTGGATGACACCTGACA 379  
QY 675 TGCAGATGATTTTTTACTGAATCTGGTAGATGGTCTTACACAAATGTTGGCTGTTG 734  
Db 380 TTGTTGACGATTTCTACTCTCACTTGTCTGACTGGGGAAGTGCAAATGCTTAGCCATAG 439  
QY 735 GTTTGGGTAATGTTGCTATCTCTGGAATGCTTGTAGCAGCAAGGTAATTAATATG 794  
Db 440 CGTTGGACCACTGCTCTCTGTTGGGATGCTTCCACTGTTCTACATCTGAGCTTGTGA 499

Search completed: January 23, 2004, 23:19:37

Job time : 721 secs

QY 795 ATTTGGGGTGTGATGATG---TGTTTGTCTGTGTGGGCTCAACGTTGGTACTCATC 851  
Db 500 CCATTGATGAGGAGAAAGGACCTGTCAAGATGATCAACTGGGCTCTGATGGTCTCATG 559  
QY 852 TTGCTGTTGGNACTAAACAATGGTAAAGTTTCAAGTTTGGGA---TGCAAGCAAGATGCAAGA 908  
Db 560 TTGCAGTTGGACTCAACAATCTGTAAGTCCAGCTGTGGGATTTGCAATCAACGTTCAAC 619  
QY 909 AGATAAGATCAATGGAGGGCCATCGGTTAGCTGTGCGGGGCTTGGCCCTGGAGTTTCACTC 968  
Db 620 TGAGAACATTAAAGGGTGGTCAACAGTCAAGAGTAGGATCACTGGCATGGAACAATCAT 679  
QY 969 TTTTGTCTTCTGGTGGAGGGAATAAGATATTTATCAAGAGATATACGACATACCAAGAAG 1028  
Db 680 TCCTTACTACTGGAGGAATGGATGGAATGATCATCAACAATGATGTGAGGATCAGATCAC 739  
QY 1029 ATTTTGTAGTAAACTGTCAAGACACAAATCAGAGGTTTGGGACTGAAGTGGTCAATG 1088  
Db 740 CCATTGTGGAACCTTACAGAGGTCACACTCAAGAAGTTTGTGGCTCAAGTGGTCAAGAT 799  
QY 1089 ATAAACGTTGATTTGGCATCTGGAGGAAATGACAAATTTGTTTGTGTTGGAATC----- 1142  
Db 800 CTGACACAACTAGCAAGTGTGGCAACGCAATGTGGTACACATCTGGGATCGTTCTG 859  
QY 1143 -----AACACTCAACCCAGCCTGCTCCTCAAGTACTGTGAGCACAACAGCAGCTG 1190  
Db 860 TCGCTTCTCAAACTCAACACACAAATGGCTGCACAGCTTGAGGACACATACATCTGCTG 919  
QY 1191 TTAAGCTATTGATGGTCTCTCATCTTCAAGACTTTCTTGCATCTGGAGGAGAACTG 1250  
Db 920 TGAAGGCTCTTGGCTGGTGGCTTTTCCAAGCGAAATTTGCTTGCACCTGGTGGTGGAG 979  
QY 1251 CAGATAGATGATTTGTTTTTGGAAATACAAACCAAACTCAACCTTAGCTGTATGGACA 1310  
Db 980 GAGACAGGACGATAAGTTCTGGAATACTCACTGGGCTTGTGTAATTCAGTAGACA 1039  
QY 1311 CTGGAAGTCAAGTTTGCATCTTCTGTGGTCCAAAATGTCACGAACTAGTAAGCACAC 1370  
Db 1040 CTGGTTCCCAAGTTTGTCTGTGTTATGGAGCAAGAAATGAAAGAGAGTTGCTTAGCTCAC 1099  
QY 1371 ATGGGTACTCCAGAACCAAGATTTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGA 1430  
Db 1100 ATGGGTTTACACAGAACTCAGCTTACACTTTGGAAGTACCCATCCATGGTGAATGGCTG 1159  
QY 1431 CTCTTACCGGCCATCTATAGGTTCTCTATCTTGGCATCTCTCCAGATGGACAGACTA 1490  
Db 1160 AGCTCACTGGTCTACATCAAGAGTTCTATATATGCGCCAGAGTCCAGATGGTTGACTG 1219  
QY 1491 TTGTAAGTCAAGCTGGAGTGAACGCTTAGGTTCTGGATGTTTT 1536  
Db 1220 TAGCTTACGAGCAGGAGACGAGACCTTAGGTTTGGAACTGTTTT 1265

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 11:32:03 ; Search time 4356 Seconds

(without alignments)

11192.557 Million cell updates/sec

Title: US-09-701-572-1

Perfect score: 2006

Sequence: 1 gattcgccagcaggaagaa.....agaaaaaaaaaaaaaaaaaaaaa 2006

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_eston:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	667.2	33.3	759	10	BG584091
2	624.6	31.1	820	14	CA919262
3	620.6	30.9	630	12	BI270288
4	575.4	28.7	594	13	BQ146406

5	571.8	28.5	1633	11	AY112458
6	543.6	27.1	648	10	BG044933
7	499.4	24.9	677	13	BU042571
8	431.8	24.5	558	10	AW980816
9	468.2	23.3	987	10	BG444274
10	463.4	23.1	528	14	CB828401
11	456.8	22.8	680	14	CA837263
12	407	20.3	856	14	CB675428
13	392	19.5	710	9	AW030735
14	376.4	18.8	460	9	AW498616
15	342	17.0	909	14	CA764180
16	339.4	16.9	501	9	AI895812
17	331.2	16.5	1818	11	AK081476
18	327	16.3	577	14	CA020821
19	312.8	15.6	533	9	AW065517
20	311.8	15.5	2982	11	AK083686
21	303.6	15.1	350	9	AU252139
22	292.8	14.6	530	13	BQ754977
23	282.6	14.1	785	12	BI652526
24	281.2	14.0	557	10	BE329969
25	281.2	14.0	705	14	CD003448
26	281	14.0	913	12	BI904176
27	280.6	14.0	554	9	AI770538
28	278.8	13.9	570	14	CA638695
29	275	13.7	731	14	CB979301
30	273.6	13.6	873	12	BI153551
31	273.4	13.6	801	11	AY109717
32	273.2	13.6	591	10	BF003530
33	273.2	13.6	878	10	BE456106
34	270.6	13.5	571	12	BI420244
35	270.4	13.5	879	13	BQ962474
36	270.2	13.5	863	12	BI554814
37	269.2	13.4	614	14	CD003026
38	267.4	13.3	688	14	CA923074
39	265.6	13.2	469	12	BM500135
40	264.6	13.2	779	13	BQ572197
41	263.8	13.2	872	12	BI456001
42	263.2	13.1	647	14	CA925058
43	263.2	13.1	647	14	CA925755
44	262.2	13.1	848	12	BM945914
45	261.4	13.0	730	14	CB979227

ALIGNMENTS

RESULT 1  
BG584091  
LOCUS  
DEFINITION  
BG584091 MHAM Medicago truncatula/Glomus versiforme mixed EST  
Library cdna clone pMEAM-14B17 5' end, mRNA sequence.  
759 bp mRNA linear EST 11-APR-2001  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
Harrison, M.J., Liu, J., Town, C.D., Van Aken, S., Utterback, T., Cho, J.  
and Frazer, C.M.  
ESTs from roots of Medicago truncatula after colonization with  
Glomus versiforme, 2001  
Unpublished  
Contact: Harrison M.J.  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73401  
Tel: 580-223-5810  
Fax: 580-221-7380  
Email: mharrison@noble.org  
Noble EST name: N379561e TIGR sequence name: WTDB090TK More  
information is available at: http://www.medicago.org  
Seq primer: SKnod (CTA GAA CTA gtg gat CC).

AY112458 Zea mays  
BG044933 saas3d02.  
BU042571 PP LRA001  
AW980816 EST391969  
BG444274 GA\_Ba002  
CB828401 L1NEST86b  
CA837263 MTU003C07  
CB675428 OBNRe11C  
AW030735 EST273990  
AW498616 AJ498616  
CA764180 AF53-Rpf  
AI895812 EST265255-  
AK081476 Mus muscu  
CA020821 HZ37018r  
AW065517 614059E06  
AK083686 Mus muscu  
AU252139 AU252139  
BQ754977 EBed02\_SQ  
BI652526 603299431  
BE329969 sc071g10.Y  
CD003448 EST0169.N  
BI904176 60316804  
AI770538 606054A05  
CA638695 wreln.pk0  
CB979301 CAB40007  
BI153551 602917131  
AY109717 Zea mays  
BF003530 EST432028  
BE456106 FVMSB001  
BI420244 L1NEST549  
BQ962474 AGENCOURT  
BI554814 603236792  
CD003026 EST1084.N  
CA923074 EST640792  
BM500135 PAC000000  
BQ572197 UT-M-PB0  
BI456001 603170445  
CA925058 MTU7TL.P1  
CA925755 MTU7TL.P5  
BM945914 UT-M-EM0-  
CB979227 CAB40007 -

```

FEATURES
  source      Location/Qualifiers
  i. .759     /organism="Medicago truncatula/Glommus versiforme mixed EST
              library"
  /mol_type="mRNA"
  /cultivar="Medicago truncatula genotype A17"
  /db_xref="taxon:119092"
  /clone="pVHAM-14817"
  /tissue_type="roots colonized with Glommus versiforme"
  /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
  post-inoculation with Glommus versiforme. The library was
  made from a mixture of RNA from each of these stages."
  /lab_host="E. coli strain XLOLR"
  /clone_lib="MHAM"
  /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
  XhoI; cDNA was prepared from polyA+ enriched RNA from
  roots harvested at 10, 17, 22, 31 and 38 days
  post-inoculation with Glommus versiforme. The cDNA was
  directionally ligated into the Unizap XR vector from
  Stratagene and packaged using Gigapack III Gold packaging
  extracts. Plasmids containing cDNA inserts were excised
  from the recombinant lambda-Zap phage using Ex-assist
  helper phage and propagated in XLOLR cells."
  BASE COUNT      187 a 165 c 185 g 222 t
  ORIGIN
    Query Match      33.3%; Score 667.2; DB 10; Length 759;
    Best Local Similarity 95.1%; Pred. No. 3.1e-78;
    Matches 721; Conservative 0; Mismatches 33; Indels 4; Gaps 3;

Qy 109 ATATAAAGCGCTCAAAAATCTTTTACAGCGTCTCTTTTCCCGGGGAAAAAATTA 168
Db 1 ATATAAAGCGCTCAAAAATCTTTTACAGCGTCTCTTTTCCCGGGGAAAAAATTAAC 60

Qy 169 ACACAGCTCCGCCATGACGGAACCGGTAATCGAAATCCACACCGACTTCCACCGTCAG 228
Db 61 ACA-AGCTCCGCCATGACGGAACCGGTAATCGAAATCCACACCGACTTCCACCGTCGG 119

Qy 229 AGACAAATCTCACCGCTGAGCACAACCGGAGAGTCTCGGTCATCTAGACCGGTATGAT 288
Db 120 AGATAAATCTCCACCGCGGAGCCATCACCGGAGAGTCTCGGTCATCTAGACCGGTATGAT 179

Qy 289 CAACAGCAACCATATACACTCTCTCGAAACCAATCTACTCCGATAGGTTCAATCCGAG 348
Db 180 CAACAGCAACCATATACACTCTCTCGAAACCAATCTACTCCGATAGGTTCAATCCGAG 239

Qy 349 TAGATCTGCTTCGAAATTCGCTTTGTTGATATCAATCTCCGACAGAGGACCGATGA 408
Db 240 TAGATCTGCTTCGAAATTCGCTTTGTTGATATCAATCTCCGACAGAGGACCGATGA 299

Qy 409 TAGTTCACGCGTTATACGACTCTCTCGAGAACCGCGTTGTTGACCGGATGTTGCCGG 468
Db 300 TAGTTCACGCGTTATACGACTCTCTCGAGAACCGCGTTGTTGACCGGATGTTGCCAGG 359

Qy 469 TCCGGTTACCGCGGAAAAACCGACTCCCGTTCGATGACATTCGCGAATAGGAATATTTT 528
Db 360 TCCGGTTACCGCGGAAAAACCGACTCCCGTTCGATGACATTCGCGAATAGGAATATTTT 419

Qy 529 TAGGTATAGAGGAGGAGAGACAGTCCATGCACTCGCTTTTCGCGTTTATGGATGATGA 588
Db 420 TAGGTATAGAGGAGGAGAGACAGTCCATGCACTCGCTTTTCGCGTTTATGGATGATGA 479

Qy 589 TTTTGTTCCTGGTGTAAATCATAGTCCGGTTAAGGCTCCTAGGAAGGTTCTTCGATGCC 648
Db 480 TTTTGTTCCTGGTGTAAATCATAGTCCGGTTAAGGCTCCTAGGAAGGTTCTTCGATGCC 539

Qy 649 TTATAAGGTTTGGATGCACCTGCTTCGACATGATTTTATCTGAATCTGGTAGATTG 708
Db 540 TTATAAGGTTTGGATGCACCTGCTTCGACATGATTTTATCTGAATCTGGTAGATTG 599

Qy 709 GTCTTCACACATGTGTTGGCTGTTGGTTTGGGTAATCTGTGTCTATCTCTGGAATGCTG 768
Db 600 GTCTTCACACATGTGTTGGCTGTTGGTTTGGGTAATCTGTGTCTATCTCTGGAATGCTG 659

```

---

```

Qy 769 TAGCAGCAGAGTAACTAAATTTATCTGATT-TGGGGGTTGATGATTGTG--TTTGTCTGT 825
Db 660 TAGCAGCAGAGTAACTAAATTTATCTGATTGTTGGGGGTTGATGAATGGGTTTGTCTGT 719

Qy 826 TGGTGGGCTCAACGCTGCTACTCATCTGCTGTTGGAA 863
Db 720 TGGTGGGCTCAACGCGGTACTCATCTGCTGTTGGAA 757

RESULT 2
CA919262/c      820 bp      mRNA      linear      EST 27-DEC-2002
LOCUS      EST636980 MTUS Medicago truncatula cDNA clone MTUS-11B5, mRNA
DEFINITION      sequence.
ACCESSION      CA919262
VERSION        CA919262.1
KEYWORDS       GI:27406192
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 820)
VandenBosch,K., Endre,G., Silverstein,K. Town,C.D., Van Aken,S.,
Utterback,T., Cheung,F. and Fraser,C.M.
The Medicago truncatula 6K unigene set: cDNA clones selected and
re-arrayed from various libraries
Unpublished
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbbs.umn.edu
TIGR sequence name: MTUAK17V
More information is available at: www.medicago.org
Seq primer: gta ata cga ctc act ata ggg c.
FEATURES
  source      Location/Qualifiers
  1..820      /organism="Medicago truncatula"
              /mol_type="mRNA"
              /cultivar="A17"
              /db_xref="taxon:3880"
              /clone="MTUS-11B5"
              /tissue_type="mixed tissues"
              /dev_stage="various stages"
              /lab_host="XLOLR"
              /clone_lib="MTUS"
              /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
              XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
              was directionally ligated into the Unizap XR vector from
              Stratagene and packaged using Gigapack III Gold packaging
              extracts. Plasmids containing cDNA inserts were excised
              from the recombinant lambda-Zap phage using Ex-assist
              helper phage and propagated in XLOLR cells."
  BASE COUNT      271 a 155 c 161 g 233 t
  ORIGIN
    Query Match      31.1%; Score 624.6; DB 14; Length 820;
    Best Local Similarity 94.9%; Pred. No. 1.1e-72;
    Matches 701; Conservative 0; Mismatches 29; Indels 9; Gaps 5;

Qy 1272 GGAATACACCAAACTCACACTTAGCTTAGTATGACATCGAGTTCAGGTTGCAATC 1331
Db 820 GGAATACACCAAACTCACACTTAGCTTAGTATGACATCGAGTTCAGGTTGCAATC 762

Qy 1332 TTTCTCTGGTCCAAAATGTCAACGAATAGTAGACACATCGGTACTCCCAAGAACCA 1391
Db 761 TTTCTCTGGTCCAAAATGTCAACGAATAGTAGACACATCGGTACTCCCAAGAACCA 702

```



```

Db      601 GAATGACACAAATTGTTTGGGAATC 630

RESULT 4
B0146406
LOCUS   B0146406
DEFINITION
NF069G04FL1F1036 Developing flower Medicago truncatula cDNA clone
ACCESSION
B0146406
VERSION
B0146406.1 GI:20283465
KEYWORDS
EST.
ORGANISM
Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 594)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
JOURNAL
Unpublished
COMMENT
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 594 Std Error: 0.00
Plate: 069 row: G column: 04
Seq primer: TCACACGGAACAGCTATGAC.
Location/Qualifiers
1..594
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF069G04FL"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/clone_lib="Developing flower"
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
Xli-Blue MRP (Stratagene). Excised plasmids were plated
using SOLR cells."
BASE COUNT      141 a 144 c 141 g 167 t 1 others
ORIGIN

Query Match      28.7%; Score 575.4; DB 13; Length 594;
Best Local Similarity 98.0%; Fred. No. 3.8e-66;
Matches 582; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      170 CACAGTCGCGCATGGAGGACCGGTAATCGAATCCACCACCGACTCCACCGTCGAGA 229
Db      1 CACAGTCGCGCATGGAGGACCGGTAATCGAATCCACCACCGACTCCACCGTCGAGA 60
QY      230 GACAAATTCACCGCGCTGAGGCATCAGCGGAGCTCCGTCATGTAAGCGGTATGATC 289
Db      61 GATAATTCACCGCGCGGAGGCATCAGCGGAGCTCCGCGCATGTAAGCGGTATGATC 120
QY      290 AACAGCAACCAATACACCTCACCTTCTCGAACAATCTACTCCGATAGTTCATTCCGAGT 349
Db      121 AACAGCAACCAATACACCTCACCTTCTCGAACAATCTACTCCGATAGTTCATTCCGAGT 180

350 AGATCTGCTTCGAATTCGCTTTGTTGATATCACTCCGACAGAGGACGCGATGAT 409
181 AGATCTGCTTCGAATTCGCTTTGTTGATATCACTCCGACAGAGGACGCGATGAT 240
410 AGTTCCAGCGCTTATACGACTCTTCTGAGAACGGCGTTGTTGGACCGGATGTCGGGT 469
241 AGTTCAGCGCTTATACGACTCTTCTGAGAACGGCGTTGTTGGACCGGATGTCGGGT 300
470 CGGTTACGCGGAAAACCGACTCCGCGTCGATGACATTCGCGAATAGGAATATTTT 529
301 CCGGTTACGCGGAAAACCGACTCCGCGTCGATGACATTCGCGAATAGGAATATTTT 360
530 AGGTATAAGACGAGACGAGACAGTCCATGCACTCGCTTTGCGCGTTTATGATGATGAT 589
361 AGGTATAAGACGAGACGAGACAGTCCATGCACTCGCTTTGCGCGTTTATGATGATGAT 420
590 TTGTTCTGCTGGTGAATCATAGTCCGTTAAGGTCCTAGGAGGTTCTTCGATCGGCT 649
421 TTGTTCTGCTGGTGAATCATAGTCCGTTAAGGTCCTAGGAGGTTCTTCGATCGGCT 480
650 TATAAGGTTTTCGATGCACTGCTTTGCAAGATGATTTTATCTCAATCTGTTAGATTGG 709
481 TATAAGGTTTTCGATGCACTGCTTTGCAAGATGATTTTATCTCAATCTGTTAGATTGG 540
710 TCTTACACAAATGTTGGCTGTTGGTTGGTAACTGCTGCTATCTCTGGAAT 763
541 TCTTACACAAATGTTGGCTGTTGGTTGGTAACTGCTGCTATCTCTGGAAT 594

RESULT 5
AY112458 1633 bp' mRNA linear HTC 17-OCT-2002
LOCUS   AY112458
DEFINITION
Zea mays CL250.1 mRNA sequence.
ACCESSION
AY112458
VERSION
AY112458.1 GI:21217048
KEYWORDS
HTC.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 1633)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1633)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1..1633
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:631015"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize

```



```

BASE COUNT      389 a      356 c      386 g      391 t      111 others
ORIGIN

Query Match      28.5%; Score 571.8; DB 11; Length 1633;
Best Local Similarity 72.8%; Pred. No. 5.7e-66;
Matches 724; Conservative 0; Mismatches 264; Indels 6; Gaps 1;

QY 622 GGCTCCCTAGGAAGTTCCTCGATCGCTTATAAGTTTTCGATGCACTGCTGTTGCAAGA 681
DB 31 GGGGCCAAGAGAGTCCCGAGGTCGCCCTCAAGGTGCTGATGCGCTGCGTTGGAGA 90
QY 682 TGATTTTATCTGAATCTGTPAGATGTTCTTACCAATGTTGCTGCTGTTGGTGG 741
DB 91 CGACTTCTACTCAACCTCGTCGATGTTCTTACATATAGCTCTGCTGTTGGTGG 150
QY 742 TACTGTGCTATCTCTGGAATGCTTGTACGACGAGTAACATAATATGATGTTGG 801
DB 151 CAATGCGTTTACTCTGTTGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 210
QY 802 GGTGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 861
DB 211 NNNNNNNNCAACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 270
QY 862 AACTAACTAAGTGAAGTTCAGATTTGGGATGCGACAGATGCAAGATATAGATCAAT 921
DB 271 AACAAATCAAGAAAGTTCAGATCTGGGATGCAACTCGCTGTAAAAGATAAGAACTAT 330
QY 922 GGAGGGCATCGTGTAGCTGTGCGGGCTTGGCTGAGTTGATCTCTTTTGTCTCTGG 981
DB 331 GGAAAGCCATCGCATGCTGTAGGTGCTCTTGATGAGTCTTCTGTTCTTCTG 390
QY 982 TGGACGGATAGAAATATTTATCAACAGATATAGCACACAGAGATTTGTTAGTAA 1041
DB 391 AAGCGTGCACAGAGATCTCCACCATGATATCCGCTCAAGAGATCATGTTAGCAA 450
QY 1042 ACTGTCAAGACACAAATCAGAGTTTGTGACTGAAGTGTCTATATATCAACCGTGAGTT 1101
DB 451 GCTTACTGGGATATAAATCTGAGTTTGTGACTCAAGTGTCTTATGACAACTGTCAGCT 510
QY 1102 GGATCTGGAGAAATGACAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1161
DB 511 TGCATCGGTTGCAATGACACAGACTTTTGTGTTTGGAAATCCATTCAGTACAGCCAGT 570
QY 1162 CCTCAAGTACTGTGAGCACACAGAGCTGTTTAAAGTATTTGATGTTCTCTCATCTTCA 1221
DB 571 ACTGAAGTATCTGAGCACACAGAGCTGTCAAGAGTATTTGCTGTTCACTCATCTCCA 630
QY 1222 TGGACTTCTTGCATCTGAGGAGAACTGACAGATAGATGTTATGTTTGGAAATACAC 1281
DB 631 TGGGCTGCTTGCATCTGTTGAGGAACTGACAGATAGATGATACGATTTTGGAAATACGAC 690
QY 1282 CACAACTCACACTTACTGTATGACACTGGAAGTTCAGTTTGCATCTTGTCTGCTC 1341
DB 691 CACAAATGCACATTTGACTGTGAGACCTGGCAGCCAGCTTGCATCTTGTGCTGCTC 750
QY 1342 CAAAAATGTCACAACTAGTAAAGCACACATGSGTACTCCAGAACACAGATTTATGTTTG 1401
DB 751 AAAGAATGTAATGAGCTTGTAGCACCCAGGCTACTCTCAGAACCAATAATTTGTTG 810
QY 1402 GAGATACCCATATGTAAGCTGGGACTCTTACCGCCACTTATATAGGTTCTCTA 1461
DB 811 GAGATATCCAAATGTCGAAGCTTGCACCTTACAGGGGCATATACATACAGATTTGTA 870
QY 1462 TCTTCCCATCTCTCCAGATGACACACTATTTGTAATCTGGAGCTGGAGATGAAGCGTTAG 1521
DB 871 CTTAGCTATTTCCCTGATGCCAGACCATAGTTACTGTTGCTGTGATGAACGCTCCG 930
QY 1522 GTTCTGGAATGTTTCCCTTCCCTTAAATCTACAGATCTCAAAAGTGAAA-----TCGG 1575
DB 931 GTTTTGGAAACGTTTCCCTCTCCAAAGTCCAGAGTTCTCAGAGTTTAAAGTTCGGTGG 990
QY 1576 AGCATTTATCTTCTGGAAGAACTACTATCATCAGTGTA 1609

Db 991 AGGAACATCTTTTCGTTAGAGCTACATCCGGTGA 1024

RESULT 6
BG044933
LOCUS
DEFINITION
    ssa35d02.y1 Gm-cl059 Glycine max cDNA clone GENOME SYSTEMS CLONE
    ID: Gm-cl059-939 5' similar to TR:Q9XED5 Q9XED5 CELL CYCLE SWITCH
    PROTEIN. :, mRNA sequence.
ACCESSION
    BG044933
VERSION
    BG044933.1 GI:12492159
SOURCE
    Glycine max (soybean)
ORGANISM
    Glycine max (soybean)
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
    ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
    1 (bases 1 to 648)
    Shoemaker R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
    A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
    Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
    Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
    R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
    R., Waterston, R. and Wilson, R.
    Public Soybean EST Project
    Unpublished
    Contact: Shoemaker R/Public Soybean EST Project
    Public Soybean EST Project
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: eschwats@wustl.edu
    This clone is available through: ResGen, Invitrogen Corp. 2130
    South Memorial Parkway Huntsville, AL 35801 For further information
    call: (800)-533-4363 or contact via email: ccu@resgen.com
    High quality sequence stop: 405.
FEATURES
    Location/Qualifiers
    1..648
    /organism="Glycine max"
    /mol_type="mRNA"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl059-939"
    /tissue_type="whole seedling, 2 week old, etiolated"
    /lab_host="DH10B"
    /clone_lib="Gm-cl059"
    /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
    XhoI; The cDNA library was constructed from mRNA isolated
    from 2 week old etiolated whole seedlings of PI468916.
    Complementary DNA was synthesized from mRNA using a primer
    consisting of a poly(dT) sequence with a XhoI restriction
    site. EcoRI adapters were ligated to the blunt-ended cDNA
    fragments followed by XhoI digestion. The cDNA fragments
    were directionally cloned into the EcoRI-XhoI restriction
    site of the pBluescript vector. The ligated cDNA fragments
    were transformed into DH10B host cells (Gibco BRL). This
    library was constructed in the laboratory of Dr. Randy
    Shoemaker at Iowa state university."
    BASE COUNT      173 a      132 c      161 g      181 t
    ORIGIN

Query Match      27.1%; Score 543.6; DB 10; Length 648;
Best Local Similarity 90.0%; Pred. No. 5.2e-62;
Matches 582; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 860 GGAACCTAACATGGTAAAGTTCAGATTTGGGATGCGAAGATGCAAGAGATAGATCA 919
DB 1 GGAACCTAGCATGGTAAAGTTCAGATTTGGGATGCGATCTCGATGCAAGAGATAGATCT 60
QY 920 ATGGAGGGCCATCGGTTACGTTGCGGGCGCTTGGCGCTGAGTTGATCTCTTTGCTTCT 979
```



```

VERSION
KEYWORDS
SOURCE
ORGANISM

AW980816.1 GI:8172357
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 558)
AUTHORS
Pedrova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,J.S., Peng
,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.
TITLE
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula
JOURNAL
Unpublished
COMMENT
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name:M256650e
TIGR sequence name:MTCB29TK
More information is available at: http://chrysis.tamu.edu/medicago
Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
FEATURES
source
1..558
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-58P10"
/tissue="N2-fixing root nodules"
/dev_stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XL0LR"
/clone_lib="GVN"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XL0LR cells."
BASE COUNT
172 a 112 c 120 g 154 t
ORIGIN
Query Match 24.5%; Score 491.8; DB 10; Length 558;
Best Local Similarity 93.3%; Pred. No. 3.4e-55;
Matches 514; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1260 GTATTGTTTGGAAATACACCAACCTACACCTTAGCTGTATGGACACTGGAAGTC 1319
Db 1 GTATTGTTTGGAAATACACCAACCTACACCTTAGCTGTATGGACACTGGAAGTC 60
QY 1320 AGTTTGCATCTTGTCTGGTCCAAAATGTCAACGAACCTAGTAAGCACACATGGGTACT 1379
Db 61 AGTTTGCATCTTGTCTGGTCCAAAATGTGGATGAACCTAGTAAGCACACATGGGTACT 120
QY 1380 CCCAGACCCAGATTAATGTTTGGAGATACCCCACTATGTCAAGCTGGCGACTTTACCG 1439
Db 121 CCCAGACCCAGATTAATGTTTGGAGATACCCCACTATGTCAAGCTGGCGACTTTACCG 180
QY 1440 GCCATACCTATAGGGTCTCTATCTTGCCATCTCTCCAGATGGACAGACTATTGTAACTG 1499
Db 181 GCCATACCTATAGGGTCTCTATCTTGCCATCTCTCCAGATGGACAGACTATTGTAACTG 240
QY 1500 GAGCTGAGATGAACCGTTCAGGTTCGGAAATGTTTCCCTCCCTTAATCAAGATA 1559
Db 241 GAGCTGAGATGAACCGTTCAGGTTCGGAAATGTTTCCCTCCCTTAATCAAGATA 300

1560 CTGAAAGTGAATCGGAGCATTATCTCTTGGAGAACTACTATCAGGTGATGATCTCTGG 1619
301 CTGAAAGTGAATCGGAGCATTATCTCTTGGAGAACTACTATCAGGTGATGATCTCTGG 360
1620 CGTTGCAGCCCAATCATGTGTCATATTTCTAAGTTTGGTTGCTGTGTAGAACTAAATTT 1679
361 CATTCGACCCCTATCATGTGTCATATTTCTAAGATTGGAGAGATAGTAGTAAATTTA 420
1680 CTGACGGGAGAACACCATGTGTGGAACAACTTGAATATAAAACACCAACCAAGTAGCAT 1739
421 TATAGCATAGAACACCATGTGAGGATATAACCTAGAGTATATAAAACACCAACCAAGTAGCAT 480
1740 CTTTACCAACTGGGAGAGCCCTTGGGAGGAGCTATAAAGTTTTCATATGCTGCCGGTGA 1799
481 CTTTACCAACTGGGAGAGCCCTTGGGAGGAGCTATAAAGTTTTCATATGCTGCCGGTGA 540
1800 TATTCCTGCAT 1810
541 TATACCTTTTAT 551

RESULT 9
BG444274
LOCUS
DEFINITION
BG444274
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gossypium arboreum
Gossypium arboreum
Eukaryota; Viridiplantae;
Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids 1; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
1 (bases 1 to 987)
AUTHORS
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL
Unpublished
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence start: 40
High quality sequence stop: 800.
FEATURES
source
1..987
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0023N11f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT
258 a 203 c 256 g 265 t
ORIGIN
Query Match 23.3%; Score 468.2; DB 10; Length 987;
Best Local Similarity 82.2%; Pred. No. 2.9e-52;
Matches 550; Conservative 0; Mismatches 118; Indels 1; Gaps 1;
QY 939 GTGTGGGCGCTTGGCCCTGGAGTTCATCTCTTGTCTCTCTCTGTCGACGAGTAAGATA 998

```



AUTHORS Cushman, J.C.  
 TITLE An expressed sequence tag database for the common ice plant,  
 Mesembryanthemum crystallinum  
 JOURNAL Unpublished  
 COMMENT Contact: Cushman JC  
 Department of Biochemistry  
 University of Nevada  
 MS200, Reno, NV 89557-0014, USA  
 Tel: 775-784-1918  
 Fax: 775-784-1650  
 Email: jcushman@unr.edu  
 PCR Primers  
 FORWARD: T3 20mer  
 BACKWARD: T7 21mer  
 Plate: 003 row: C column: 07  
 Seq primer: T3 20mer  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1. .680  
 /organism="Mesembryanthemum crystallinum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3544"  
 /clone="MCT003C07"  
 /tissue\_type="leaf"  
 /dev\_stage="five-week-old"  
 /clone\_lib="Ice plant Lambda Uni-Zap XR expression library  
 , 5 days 0.5 M NaCl treatment, Crassulacean acid  
 metabolism, phase IV (5:30 PM)."  
 /notes="Vector: Lambda Uni-Zap XR, Bluescript SK; Site 1:  
 EcoRI; Site 2: XhoI; Library construction was performed  
 according to Stratagene's recommended protocol for the  
 Lambda UniZapXR vector and cDNA synthesis kit."  
 BASE COUNT 181 a 145 c 163 g 191 t  
 ORIGIN  
 Query Match 22.8%; Score 456.8; DB 14; Length 680;  
 Best Local Similarity 81.3%; Pred. No. 1.1e-50;  
 Matches 530; Conservative 0; Mismatches 122; Indels 0; Gaps 0;  
 QY 925 GGGCCATCGGTACGTGTCGGGCGCTTGGCTCGAGTTCATCTCTTTGCTCTCTGCTG 984  
 DB 22 GAGCATCGGTTGGGTGGGTCTTGGCATGGAGTCTCTGCTGTTGTCATCTGTTAG 81  
 QY 985 ACGGGATPAGATATTATCAACAGAGATATACGCACACAGAGATTTTGTAGTAACT 1044  
 DB 82 CCGAGATAAGATATATTCCAAAGAGATATACGTTGTGAGATGATTATGTCAGTAAAT 141  
 QY 1045 GTCAGGACAAATCAGAGGTTTGTGACTGAGTGTGTCATATGATAACCGTGAGTTGGC 1104  
 DB 142 AACAGGACAAATCAGAGGTTGTTGGCTGGAAGTGTCTTATGATAACCGTGAGTAGC 201  
 QY 1105 ATCTGGAGGAATGACAACAATTTGTTGTTGGAAATCAACACTCAACCCAGCGCTTCCT 1164  
 DB 202 ATCAGGAGGAATGACAAATAGGCTTTTGTGTGGAAACCAATCTTACTCAACAGTGT 261  
 QY 1165 CAAGTACTGTGAGCACACAGAGCTGTTAAAGCTATTGCACTGTCTCTCTCATCTTCATG 1224  
 DB 262 GAAGTACTGGGAACATACAGAGCGTGTAAAGCAATTCGTTGTGCGCCCATCTTCATGG 321  
 QY 1225 ACTTCTTGTGATCTGGAGGAGAACTGCAGATAGATGATTTCGTTTGTGGAATCAACACCAC 1284  
 DB 322 ACTCTCGCATCTGGTGGTGTACTGCTGATCGATGATCCGTTTCTGGAATCAACACCAC 381  
 QY 1285 AAATCAACCTTAGCTGTATGGACACTGGAAGTCAGTTTGCATCTTGTCTGTTGCCAA 1344  
 DB 382 TAATCTCAGTTGGGATGCACTGGACACTGGCAGTCAGTTTGCATCTTGTGGTCTAA 441  
 QY 1345 AAATGTCAACGAATAGTAAAGCACACATGGTACTCCAGAACACAGATATTGTTTGGAG 1404  
 DB 442 AAATGTGAATGAATTTGTGCACACCCATGGATCTCTCAAAACCAATCATCTTTGGAG 501  
 QY 1405 ATACCCCACTATGTCAAAGCTGGGCACTCTTACCGGCCATCTTATAGGCTTCTCTATCT 1464

Db 502 ATAGCCTAGCATGTCAAAGTTGGCAACTCTTACAGGACACACTTACAGAGTTCTTTATCT 561  
 QY 1465 TGCCATCTTCCAGATGGACAGACTATTGTAACTGGAGCTGGAGATGAACCGTTAGTT 1524  
 Db 562 AGCTATATCACCGGATGGCCAGACCATTTGTAACTGGAGCAGGAGATGAACACTGAGATT 621  
 QY 1525 CTGGAATGTTTTCCTTCCCTAAATCACAGAATATCTGAAAGTGAATCGGA 1576  
 Db 622 CTGGAATGTTTTCCTTCCCTAAATCACAGAATATCTGAAAGTGAATCGGA 673  
 RESULT 12  
 CB675428 856 bp mRNA linear EST 09-APR-2003  
 LOCUS OSJNE11C18 f OSJNEe Oryza sativa (japonica cultivar-group) cDNA  
 DEFINITION Clone OSJNE11C18 5', mRNA sequence.  
 ACCESSION CB675428  
 VERSION CB675428  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 856)  
 AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 TITLE Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 JOURNAL Unpublished  
 COMMENT Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR Primers  
 FORWARD: gta aaa cga cgg cca gtg  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 11 row: C column: 18  
 Seq primer: gta aaa cga cgg cca gtg.  
 FEATURES  
 source  
 1. .856  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNE11C18"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEe"  
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
 XhoI; 24 hrs after inoculation with Rice Blast (70-15)"  
 BASE COUNT 193 a 214 c 241 g 208 t  
 ORIGIN  
 Query Match 20.3%; Score 407; DB 14; Length 856;  
 Best Local Similarity 75.9%; Pred. No. 3.2e-44;  
 Matches 503; Conservative 0; Mismatches 160; Indels 0; Gaps 0;  
 QY 616 GGTTAAGCTCTCTAGGAAGTTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTT 675  
 Db 194 GGGCTAGTGGCCCAAGAGATCCCTAGTCACTTATAGGTGCTGGATGCTCCGCATT 253  
 QY 676 GCAAGATGATTTTATCTGAATCTGGTAGATGTTGCTTTCACAAATGTTGGCTGTGG 735  
 Db 254 GCAGATGACTTCTACCTGAACCTGTGGATTGTGCTTCGCAATAATATCTTCGAGTTGG 313  
 QY 736 TTGGGTAACTGTGCTATCTCTGGAATGCTTGTAGCAGCAAGGTAACTAATATATGGA 795

Db 314 ATTGGGGAATGTGCTTACTTATGGAATGATGCGACGACAAAGGTCCAGCTATGTGA 373  
Qy 796 TTTGGGGTTCATGATTTGTTGTTCTGTTGGCTCAACGTTGTTACTCATCTTGC 855  
Db 374 TTTGGGGTTCATGATTTGTTGTTCTGTTGGCTCAACGTTGTTACTCATCTTGC 433  
Qy 856 TGTGGGACATCAACATGTTGATTTGATTTGGGATGCGACGATGCAAGAGATAAG 915  
Db 434 TGTAGGACAAACCAAGGCAAGTTTCTGTTGGATGCGACCTCGATGTTAGAGATTAAG 493  
Qy 916 ATCAATGGAGGCCCATCGTTTACGTGTCGGGGCTTGGCTGAGTTTCATCTTTTGTG 975  
Db 494 AACCATGGAAGCCATCGATGCGAGTAGTGTCTTTCATGGAATTCATCTTCTTTC 553  
Qy 976 TTTGGTGGAGGATGATGATTTTATCAACGATATACGACATATACGACAAAGATTTGT 1035  
Db 554 GTCAAGCAGTGTGCAAGAGATCTCTTCCACATGATATCTCGTCCGAGATGATAT 613  
Qy 1036 TAGTAAACTGTGAGGACACAAATCAGAGGTTTGTGGCTGAGTGGTTCATATGATAACCG 1095  
Db 614 TACTAGACTTGTGGGCTAATCGGAGGACTGTGGGCTCAAGTGTCTTATGATAACCG 673  
Qy 1096 TGAATGGCATCTGGAGAAATGACAAATTTGTTTGGTAAATCAACACTCAACCCA 1155  
Db 674 TCAGCTTGCATCTGTTGATGATGCAACAGACTTATGATGGAATCAACACTCGGCGCA 733  
Qy 1156 GCCTGTCTCAAGTACTGTGAGCACACAGAGCTGTTTAAAGCTATTGCTGCTCTCTCA 1215  
Db 734 CCGGTACTGAGTATGACTGAGCATACAGGAGCTGCTCAAGCTATTGCGTGGTCACTCA 793  
Qy 1216 TCTTCATGAGCTTCTGATCTGAGGAGGACTGCGAGATGATGATTTGTTTGGAA 1275  
Db 794 TCTTCATGAGCTTCTGATCTGAGGAGGACTGCGAGATGATGATTTGTTTGGAA 853  
Qy 1276 TAC 1278  
Db 854 TAC 856

RESULT 13  
AW030735  
LOCUS EST273990 tomato callus, TAMU Lycopersicon esculentum cDNA clone  
DEFINITION cLSEC25N23, mRNA sequence.  
ACCESSION AW030735  
VERSION AW030735.1 GI:5889491  
KEYWORDS  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

FEATURES  
source  
1..710  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLSEC25N23"  
/tissue\_type="callus"  
/dev\_stage="25-40 days old"

/lab host="XLI-Blue MRP"  
/clone lib="tomato callus, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni laboratory; cLSEC - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"  
BASE COUNT 209 a 142 c 149 g 209 t 1 others  
ORIGIN

Query Match 19.5%; Score 392; DB 9; Length 710;  
Best Local Similarity 79.5%; Pred. No. 3.4e-42;  
Matches 464; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 1026 AAGATTTTGTAGTAAACTGTGAGGACAAATCAGAGGTTTGTGGACTGAAGTGGTCAT 1085  
Db 1 AAGATTTTGTAGTAAACTGTGAGGACAAATCAGAGGTTTGTGGACTGAAGTGGTCAT 60  
Qy 1086 ATGATTAACCGTGAATTTAGCTTTCAGGTGGAATGATAATCGGCTTTTGTATGGAATCAAC 1145  
Db 61 ATGATTAACCGTGAATTTAGCTTTCAGGTGGAATGATAATCGGCTTTTGTATGGAATCAAC 120  
Qy 1146 ACTCAACCCAGCTGCTCTCAAGTACTGTGAGCACACAGAGCTGTTTAAAGCTATTGTCAT 1205  
Db 121 ATTCAACCAACCTGTGCTGAAATCTGTGAGCATCTGCTGCGGTTAAGSCCATTTGCAT 180  
Qy 1206 GGTCTCTCTCATCTTCATGAGACTTCTTGCATCTGAGGAGGAACTGCGAGATGATGATTC 1265  
Db 181 GGTCTCTCTCATCTTCATGAGACTTCTTGCATCTGAGGAGGAACTGCGAGATGATGATTC 240  
Qy 1266 GTTTTGTGAATACACCAACCAACCTTACCTTGTATGAGCACTCGAAGTCAAGGTTT 1325  
Db 241 GATTTGTGAATACACCAACCAACCTTACCTTGTATGAGCACTCGAAGTCAAGGTTT 300  
Qy 1326 GCAATCTTCTGTGTCGCAAAATGTCACAGAACTAGTAAAGCACATGGGTACTCCGAGA 1385  
Db 301 GCAATCTTCTGTGTCGCAAAATGTCACAGAACTAGTAAAGCACATGGGTACTCCGAGA 360  
Qy 1386 ACCAGTATTTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCATA 1445  
Db 361 ATCAGATATATGTTTGGAGATATCCGAAATGTCGAAATGTCGAAATGTCGAAATGTCGAA 420  
Qy 1446 CTTATAGGGTCTCTTATCTTTCATCTTCCAGATCTTCCAGATGGAAGCACTTATCAGGTGA 1505  
Db 421 CATATAGAGTCTTATCTTATCTTGTCTATCTCCAGATGGAAGCACTTATCAGGTGA 480  
Qy 1506 GAGATGAACCGTCTAGGTTCTGGAATGTTTCCCTTCCCTAAATCAAGATATCTGAAA 1565  
Db 481 GAGATGAACCGTCTAGGTTCTGGAATGTTTCCCTTCCCTAAATCAAGATATCTGAAA 540  
Qy 1566 GTGAATTCGGAGCATTTATCTCTTGTGAAGAACTTACTTATCAGGTGA 1609  
Db 541 GTGAATTCGGAGCATTTATCTCTTGTGAAGAACTTACTTATCAGGTGA 584

RESULT 14  
AJ498616  
LOCUS AJ498616 MTPOSE Medicago truncatula cDNA clone mt--acc955206907,  
DEFINITION Medicago truncatula (barrel medic)  
ACCESSION AJ498616  
VERSION AJ498616.1 GI:22089059  
KEYWORDS  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula

REFERENCE  
AUTHORS  
1 (bases 1 to 460)  
Firnhaber, C., Bartelsemeier, V., Meyer, F., Bartels, D., Bekel, T., Linke, B., Puehler, A. and Kuester, H.





Db 362 ACATTGACGGCCATACATATAGGTATTATTTAGCCATCTCCCCAGATGGACAGACT 421  
Oy 1490 ATTGTAACCTGGAGCTGGAGATGAACCGCTTAGGTTCTGGAAAGTTTCCCTTCCCTTAA 1549  
Db 422 ATAGTAACCTGGCGCTGGTGTGATGAACCGCTTCGGTTTGGAAACGTGTTCCATCTCCCAAG 481  
Oy 1550 TCACAGAATACTGAAGTGAAA 1571  
Db 482 TCCAGAGTTCTGCACGCCTAA 503

Search completed: January 23, 2004, 21:05:55  
Job time : 4362 secs